8-2015

UNDERSTANDING MOLECULAR INTERACTIONS BETWEEN PROTEINS AND CARBON NANOMATERIALS

Siva Dasetty
Clemson University, siva.k.dasetty@gmail.com

UNDERSTANDING MOLECULAR INTERACTIONS BETWEEN PROTEINS AND CARBON NANOMATERIALS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Chemical Engineering

---

by
Siva K. Dasetty
August 2015

Accepted by:
Dr. Sapna Sarupria, Committee Chair
Dr. Mark Blenner
Dr. Rachel Getman
Dr. Ramakrishna Podila

ABSTRACT

Over the past few years there has been a rise in the number of nanomaterials engineered for a wide array of applications because of their unique properties. This rise in the development of engineered nanomaterials (ENMs) and its growing usage has also raised questions about its potential impact on the biological environment. Recent experimental studies suggest that these ENMs could be toxic due to the formation of a protein corona. Therefore, understanding the formation of a protein corona would provide some insights into the toxic behavior of ENMs. This requires understanding the interactions between proteins and ENMs.

We employ molecular dynamics simulations to explore the factors and governing forces influencing interactions between carbon nanomaterials (CNMs) and proteins. We first study the interactions between bovine serum albumin (BSA) and a set of CNMs that are of varying shape and surface chemistry. These CNMs include a single walled carbon nanotube (SWCNT), a graphene nanoribbon (GNR) and a graphene oxide nanoribbon (GONR). Our results indicate that BSA interacted with all the three CNMs and its interaction strength follows the order GNR>SWCNT~GONR. During this interaction, we found a strong correlation between the interaction energy and the number of heavy atoms of BSA near the CNM. However, there are no significant changes in the secondary structure content and all α helices are stable during this interaction in the timescales of our simulations. We have also not observed any one or two types of amino acids that are dominant during the interactions of BSA with the CNMs to identify the driving forces.

In order to determine the role of various factors such as i) aromatic residues ii) arginine iii) water and also iv) neighboring residues of an aromatic residue or arginine in the interactions of proteins with CNMs, we have designed a set of tripeptide-CNM systems. We used an advanced sampling method, umbrella-sampling method in order to determine the free energy of interaction between the tripeptides and CNM, which will enable us to quantify the contributions of different factors to the interactions. Our initial results show that the influence of the neighboring residues (Val/Leu/Thr/Ser/Gly) in the free energy of interactions for the tripeptide-graphene system where the central residue is Phe is not significant. However, barriers appear during the interactions for the larger and polar side chains. We hypothesize this is due to conformational changes that the larger side chains need to make as the tripeptide associates with the CNMs.

ACKNOWLEDGMENTS

First and foremost, I would like to sincerely thank my advisor Dr. Sapna Sarupria for providing me a wonderful opportunity to be a part of this project, which is not only intriguing but also challenging at the same time. Apart from learning some of the latest methods, I also gained a better understanding of the subject during this time. Her guidance was invaluable, as it helped me to understand some of the nuances involved between various scientific terminologies and also some common terminologies. One of the major hurdles is always not being able to apply the knowledge gained to a practical problem. However, working with her has at least helped me to learn the process of applying knowledge to a problem of interest to some extent. I would also like to thank her for always responding and providing feedback immediately.

Secondly, I would like to thank the department of chemical engineering for providing me some of the necessary resources and also a wonderful time. I would like to thank Dr. Hirt specifically, as his independent course on biopolymers was very helpful to this project. I would like to thank Dr. Getman for organizing the super group meetings as it helped me to understand some of the literature that used DFT simulations. I also would like to thank Dr. Podila from the physics department for explaining the background of this project and also providing details about some of the experiments and experimental methods.

Thirdly, I would like to thank my thesis committee Dr. Sarupria, Dr. Podila, Dr. Blenner and Dr. Getman for reviewing this thesis. I also would like to thank Brittany for reviewing and correcting some of the grammatical mistakes in this thesis and Ryan for

discussing some of the basics of molecular dynamics simulations. I am also thankful to Saptarshi for being my gym trainer and Murry for being my gym trainee. I know the training must be hard for Murry but I want him to remember, “No pain no gain”.

Finally, I would like to thank my entire family for their love and affection and always having confidence in me.

TABLE OF CONTENTS

Page

Table of Contents (Continued)

Page

LIST OF TABLES

LIST OF FIGURES

List of Figures (Continued)

List of Figures (Continued)

Figure Page

CHAPTER ONE

INTRODUCTION

The ability to design and build structures at an atomic scale is revolutionary in materials engineering. These materials called as nanomaterials have strikingly different properties compared to their larger counterparts mainly due to their extremely small dimensions. The dimensions of nanomaterials are typically in the range of 1-1000 nm. To provide an idea about nanoscale, a typical hair strand is around 60,000 nm and a red blood cell is around 7000 nm [1]. They are of the dimensions of most of the biological molecules like proteins, viruses and cellular membranes. A nanoparticle is a particle with at least one dimension smaller than 1000 nm and as small as atomic and molecular length scale.

Nanomaterials can be found in nature, like in the case of gecko feet, spider silk and a lotus leaf [2]. It is the dimensions of the hair on the foot of a gecko that gives it an ability to adhere to surfaces, allowing it to walk on walls while in the case of lotus leaf, it is the nanoscale structure that makes it hydrophobic allowing it to self clean. Recently, there has been great progress in synthesizing materials of these dimensions, creating a new class of materials called engineered nanomaterials (ENMs). ENMs have unique set of properties allowing them to be used in a wide array of disciplines. Some of the widely used ENMs are metallic nanoparticles (gold, silver, titanium dioxide), quantum dots and carbon nanomaterials (CNMs). The applications of the gold and silver nanoparticles range from synthetic products such as photo sensors, antimicrobial products to biological applications, which include drug delivery agents and biosensors [3],[4]. Titanium dioxide

and zinc oxide nanoparticles are often used in sunscreen lotions because they effectively block UV radiations and copper nanoparticles are used in some personal care products [5]. The main types of CNMs include carbon nanotubes (single and multi-walled), graphene sheets (single and multilayered), graphene oxide sheets and Buckminster fullerene (known as Bucky balls). Their applications are diverse including commercial products such as bicycle frames, badminton rackets, electronics, adhesives, textiles, aerospace products and even biological applications which include implants, biosensors, drug delivery agents and other healthcare products [1], [6], [7].

The growing development and usage of ENMs has also raised questions about their safety in the biological environment because of the same reasons that gives them their unique properties, i.e., their extremely small dimensions. Recent research [8], [9] suggests that the biological environment can be exposed to ENMs intentionally or unintentionally. Intentional exposure includes nanomaterials that are injected into the bodies as a drug delivery agent or nanoparticles used in food and cosmetic products such as sunscreen lotions. Unintentional exposure includes accidental inhalation or accidental dermal exposure of nanoparticles while they are synthesized. It is therefore imperative to understand the effects of ENMs on the biological environment and also understand their properties for developing safer and efficient ENMs. In general, in an intentional exposure, the nanomaterials that are used for carrying drugs are injected into the organism and are exposed to the blood stream as illustrated in Fig 1.1. Therefore, the efficient transport of the ENMs as a drug delivery agent to the target site depends on the biological environment it is exposed to, i.e., the blood stream. The blood stream

composes different biological molecules and can have a huge impact on the transport mechanism and on the native properties of the ENMs.

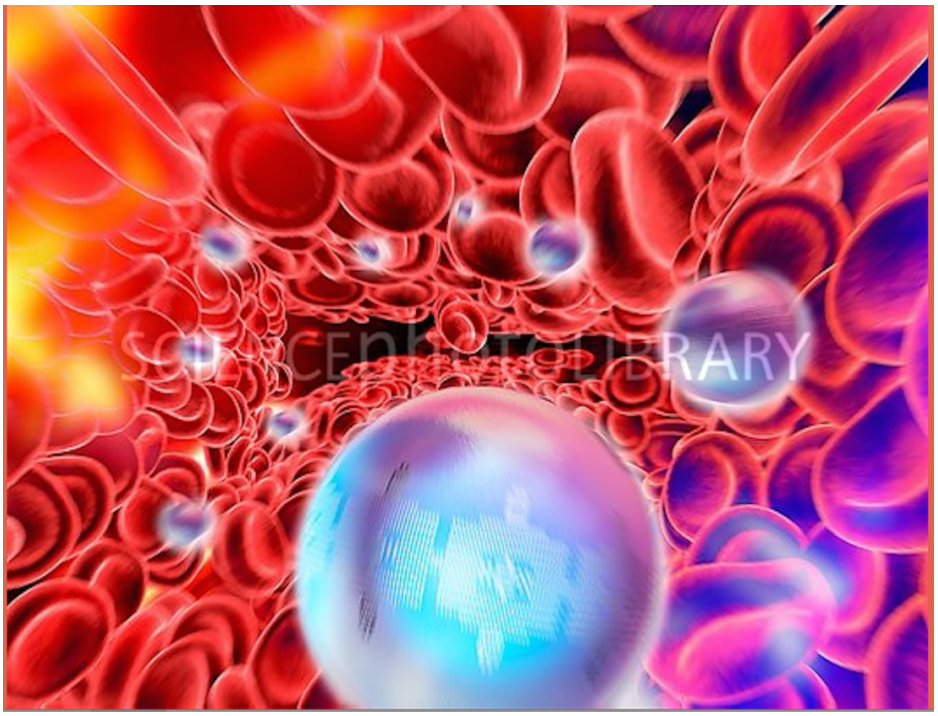


Figure 1.1: Illustration of nanoparticles in the blood stream [10].

The affinity of some of the biological molecules like proteins could indeed alter the synthetic identity and the characteristics of the ENMs [9]. Such an affinity of proteins, which results in the formation of a protein corona, could impact the conformations of the proteins. As the function of a protein depends on its three dimensional structure, the influence of the nanomaterial on its conformation could disrupt its native function and result in deadly diseases. Similarly, in the case of an unintentional exposure, inhalation of carbon nanotubes could cause asbestos-like diseases (inflammation) due to its similarity in the needle like shape. Silver nanoparticles that are

often used for its antimicrobial properties could enter the aquatic environment (present in antimicrobial clothing) [1] damaged cells derived from human and mammalian skin, lung, brain and vascular tissues in an experimental study.

The remarkable applications of nanomaterials are therefore hindered with such adverse effects. It is important to study and understand the factors responsible for such harmful effects. Once these factors are identified, understanding the interactions at a fundamental level between ENMs and proteins could help in preventing such toxic effects and assist in designing safer and efficient nanomaterials. This process of understanding toxic effects of nanomaterials, called nanotoxicity has attracted wide attention among researchers lately and is our subject of interest.

The chapters in the thesis are organized in the following manner. The second chapter includes a literature summary of the studies that explored some of the factors and the fundamental governing forces between ENMs and proteins. In the third chapter we present our work, understanding the role of shape and surface chemistry of CNMs on its interactions with the most abundant blood serum protein, albumin. The fourth chapter includes a brief introduction of another work, which is exploring the role of a neighboring residue of an aromatic amino acid in driving protein-CNMs interactions. Lastly, we summarize and propose future work in the fifth chapter. At the end, we also include appendices that show results of studies performed while implementing our simulations. Appendix A shows the performance analysis of GPUs and CPUs done prior to running our all atom albumin-CNM simulations constituting 230,000 atoms. Appendix B provides supplementary information for chapter 3. Appendix C includes a procedure

for running umbrella-sampling simulations using GROMACS. Appendix D shows the problem using a same temperature-coupling group in molecular simulations of protein-water systems and finally the last Appendix E is a brief introduction to implicit solvent simulation method.

CHAPTER TWO

LITERATURE SURVEY

This chapter provides a summary of the recent literature on the interaction of three widely used CNMs, namely carbon nanotubes (single walled and multi-walled), graphene and graphene oxide with proteins. While the focus is on the work that employs computer simulations, it also includes a perspective of recent results from *in vivo* and *in vitro* methods.

2.1 Introduction

It was not so long ago the seeds of nanotechnology were implanted and its potential applications were realized. Indeed, nanotechnology has been an industrial revolution and its global market is predicted to hit US$ 75.8 billion by 2020 [11]. It is not just the market, if one thinks over the numerous possibilities of the world at a nanoscale as pointed out in the famous lecture of Prof. Richard P. Feynman in 1959, “Plenty of room at the bottom” [12]. However, when we start exploring and utilizing the applications of the unknown world, it is important to know its implications as well. As the evidence of the toxic effects due to nanoparticles is mounting, it has become increasingly important to understand the reasons behind the functions of these nanoparticles that are not intended while implemented. One can observe the rise in the trend of the usage of the word “nanotoxicology” since 2000 in a goggle n-gram viewer [13], which provides a rough estimate about the rise in this area of research over the past few years.

Nanotoxicology is a branch of toxicology concerned with the study of the toxicity of nanomaterials, which can be divided into those derived from combustion processes (like diesel soot), manufacturing processes (such as spray drying or grinding) and naturally occurring processes (such as volcanic eruptions or atmospheric reactions) [14]. In order to prevent these adverse effects, it is important to understand the process by which these nanomaterials induce toxicity. As their larger counterparts are not toxic, one apparent reason could be related to their dimensions.

As stated in chapter 1, the exposure of nanomaterials can be categorized into a) intentional exposure and b) unintentional exposure. Both of these exposures have an equal impact on the toxic effects with the rise in the usage of products containing nanomaterials. In a 2005 study by Tsuji et al. [8], it was pointed out that the likelihood of exposure of nanomaterials is highest for workers synthesizing them or working with them. This kind of exposure is unintentional and the contact route is either by accidental inhalation or dermal exposure. Exposure to skin can also take place when sunscreen lotions that contain components like TiO_2 or ZnO nanoparticles are applied directly on the skin [5]. This is because the dimensions of an average human sweat pore size are estimated to be around 60 to 80 microns in diameter and 2 to 5 mm long [15] which are large enough to allow nanoparticles to penetrate through. Similarly, other possible routes of exposure of other nanoparticles are explained in Yah et al. [16]. Here in this literature summary, we focus only on the exposure and effects of CNMs.

Since the famous discovery of the carbon nanotube in 1991 by Dr. Sumio Iijima [17] to the recent discovery of graphene in 2004 by Sir Andre Geim [18], there have been

a number of applications associated with these nanomaterials. One of the biological applications of these nanomaterials is their usage as a drug delivery agent. In this case, the CNMs are exposed to the blood stream as illustrated in Fig 1.1. Similarly, other routes of exposure to different biological environments such as lungs can take place by accidental inhalation. This inevitable exposure of the CNMs due to their growing applications to the biological environment provides a necessity to understand their behavior and reasons for any adverse effects on the biological environment.

2.1.1 Evidence for toxic effects of CNMs

Kashiwada [19] conducted experiments to investigate the distribution of water-suspended fluorescent nanoparticles in a see-through medaka. They found that the nanoparticles are distributed throughout the body of the fish, including kidney, testis, liver, blood and even brain by crossing the blood-brain barrier. When mice are exposed to airborne nanoparticles, such as carbon nanotubes, Donaldson and Poland [20] found that carbon nanotubes get exposed to the outer lining of the lungs and cause scarring similar to asbestos. In fact, carbon nanotubes have been related to asbestos only because of its similarity in shape as stated in a recent review [21],

"The dangers of asbestos first came to light in the early 1960s, when studies linked exposure to these silicate fibers with mesothelioma – a rare cancer of the lining of the chest or abdomen that's commonly fatal. Asbestos fibers were found to be so small that they could be inhaled into the deep lung, where they could stick around for decades.

Once there, metals in the silicate fibers could act as catalysts to create reactive oxygen compounds that go on to damage DNA and other vital cellular components.
Whether nanotubes could reproduce this behavior is unknown: Their toxicity has yet to be tested. But already views on their safety differ sharply. "[Nanotubes] may be wonderful materials," says Art Langer, an asbestos expert at the City University of New York's Brooklyn College. "But they reproduce properties [in asbestos] that we consider to be biologically relevant. There is a caution light that goes on." Most notably, says Langer, nanotubes are the right size to be inhaled, their chemical stability means that they are unlikely to be broken down quickly by cells and so could persist in the body, and their needlelike shape could damage tissue."

Similarly, the other diseases due to the exposure of carbon nanotubes to various organs either by intentional or unintentional exposure are provided in detail in a review article by Madani et al. [22]. Some of the common diseases included are Alzheimer's, asthma, high blood pressure and colon cancer when nanoparticles are exposed to brain, lungs, circulatory system and gastro-intestinal system respectively. Similarly, Wang et al. [23] have performed a set of experiments on mice and showed that graphene oxide is toxic and produces lung granuloma and also may not be cleared by kidneys. In an *in vitro* study, Morales et al. [24] found that the binding of nanotubes to surfactant proteins A and D, which have a role in the lung immune defense mechanism can be damaged. In general, most of the diseases are caused due to the changes in the structure of the proteins. Therefore, it is important to know the properties of nanoparticles in a biological environment and its effects on proteins.

2.1.2 Solubilizing CNMs using proteins

Carbon nanotubes and graphene are inherently hydrophobic nanomaterials. It is therefore necessary to make them soluble in aqueous media before using them as biosensors or drug delivery agents [25]. In order to solubilize nanomaterials, different methods have been proposed. One of the methods includes chemical modification of nanomaterials. However, this modification could alter the electronic properties of the nanomaterials, which is not desired for its applications. Recently, functionalizing using non-covalent bonds by using proteins is being explored as an alternative, as this may not change the electronic properties of the nanomaterials. Ge et al. [26] have suggested that binding of proteins to carbon nanotubes can reduce the cytotoxicity and Raffaini et al. [27] suggested that proteins can be adsorbed on the inner wall of the nanotube making a lesser impact on the outer wall properties of the nanotube.

In either of the above two cases, where proteins are used to solubilize CNMs or to understand the adverse effects of CNMs, it is important to study the interactions between proteins and CNMs. Recently, there has been an increasing number of *in vivo* and *in silico* studies to address this problem and to understand the factors influencing this interaction and also the governing forces in this interaction.

2.2 Factors Influencing Protein-CNMs Interactions

As seen in the previous section, when an organism is exposed to nanomaterials, they spread out through out the organism because of their extremely small dimensions. In order to understand the possible factors of this distribution, the effects of particle size,

shape, concentration of the nanomaterials were studied experimentally and found to have varying toxic effects. Podila et al. [28] have recently summarized the physicochemical properties of the nanomaterials that influence the interactions of nanomaterials with the biological environment. They show that the size, shape, defect density, physicochemical stability and surface modification are the key properties of the nanomaterials that can influence nanotoxicity. Table 2.1 summarizes the impact of each of these key physicochemical properties.

Table 2.1: Key Properties of Nanomaterials that Influence the Biological Environment

Physicochemical Property	Impact
Size	Different groups have observed that the toxicity of different nanoparticles (Au/CdSe) was greatest when the size of the particles is around 50nm.
Shape	Compared to spherical shaped nanoparticles of Au/Ag/CdSe, their rod shaped counterparts exhibited lower toxicity. In case of nanotubes, aspect ratio (length/diameter) was found to play a key role. High aspect ratio or longer multi walled nanotubes exhibited higher toxicity when injected into subcutaneous tissue of rats.

Physicochemical stability	Inherently hydrophobic nanomaterials (Au nanoparticles, nanotubes, graphene) when functionalized by using surfactants to achieve stable suspensions changed the electronic properties of the nanomaterials and exhibited varying degrees of toxicity.
Defects and Impurities	Intrinsic defects such as stone-wales defects due to the manufacture process or extrinsic properties due to the presence of metal catalysts changed the surface reactivity and were found to produce reactive oxygen species leading to inflammation in the lungs. Similarly impurities or presence of dopants change electronic properties. Al-doped ZnO nanoparticles exhibited higher toxicity compared to pure nanoparticles and N-doped MWNT did not exhibit toxicity compared to un-doped nanotubes.
Surface modification	Non-covalent functionalization of nanomaterials such as nanotubes and graphene are found to increase their biocompatibility and exhibit lower toxicity.

Podila et al. also mention that all of the above listed properties are interrelated and have varying roles in influencing the interactions between nanomaterials and proteins,

which makes it a challenging problem to understand toxic effects of nanomaterials based on their physical and chemical properties. Nel et al. [29] proposed some of the ways to alter the designs of the nanomaterials by noting the possible mechanism of cytotoxicity, but the proposed methods could change the properties of the nanomaterials, which is not desired for the applications of nanomaterials.

Apart from the properties of the nanomaterials, the biological environment to which the nanomaterials are exposed also play a significant role on their toxic effects. Depending on the biological environment, varying types of proteins will be exposed to nanomaterials. The governing forces and the dynamics of the protein at the interface of protein-nanomaterials may depend on both the properties of the nanomaterial and the protein. Some of the factors of the proteins that influence the interaction include the primary sequence, secondary structure content, number of amino acids containing hydrophobic and aromatic side chains. The kinetics of the interaction of a protein depends on its affinity to the nanomaterial and any proteins that have a moderate affinity for the nanomaterials form a hard corona (irreversible binding) while proteins with low affinity form a soft corona (reversible binding) [28]. The governing interactions in the process of forming or after forming a protein corona are important to understand the toxic effects of nanomaterials.

2.3 Driving Forces in a Proteins-CNMs Interactions

In this section, we summarize the driving forces that are found to be dominant in many of the recent experiments (*in vivo* and *in vitro*) and computer simulations. As seen

before, there are many possible factors that influence the interactions between proteins and nanomaterials. However, the governing forces of these interactions are considered to be hydrophobic interactions, π-π stacking of aromatic side chains of proteins with the CNMs, and electrostatic forces. In most of the recent studies as highlighted in the review articles presented [30], [31], [32], the dominant interactions among these forces are found to be the hydrophobic and π-π stacking interactions for proteins and graphene/nanotubes. However, the role and significance of other factors such as water, other non-aromatic residues like arginine and a combination of both hydrophobic and hydrophilic residues in a few other studies suggest that these driving forces may not be dominant all the time. Below, we first report the studies that indicate the driving forces to be hydrophobic and π-π stacking interactions. Then we present the studies that identify the side chains of the amino acids that have the strongest interactions with CNMs and then later report the studies that indicate the role of other factors such as water, type of secondary structure, size and shape of nanomaterials.

In a combined experimental and *in silico* study by Ge et al. [26] of the four most abundant blood serum proteins and SWCNT, they have indicated the driving forces to be due to π- π stacking of side chains containing aromatic groups Tyr and Phe with SWCNT. Similarly, in the case of a graphene oxide-DNA segment system, MD simulations conducted by Chen et al. [33] have indicated the driving forces to be both electrostatic and π-π stacking interactions. Zuo et al. [34], have performed MD simulations on a protein-nanotube system and observed that a nanotube can plug into the core of the protein due to hydrophobic interactions. In a later study by the same group

[35], of a villin head piece-graphene system, where graphene was allowed to be flexible, they found that π-π stacking interactions to be the dominant driving forces. These conclusions about driving forces are also seen in some of the works that employ DFT simulations and experiments. Wang and Ai [36] have performed a set of DFT simulations using implicit solvent and MP2 methods on a tripeptide-SWCNT system, where the tripeptide is of the form G-X-G (X is any of the 20 standard amino acids). On comparing the adsorption energies of the individual amino acids in gas phase and aqueous phase (implicit solvent), they have observed that aromatic residues interact always with the nanotube. While, the neutral residues interact weakly in both the phases, the charged residues are found to have a stronger interaction only in the gas phase. In an experimental study, while identifying the peptides that have high affinity for nanotubes using phage display technique, Wang et al. [37] have found that the peptides rich in His and Trp have the highest affinity implying the dominance of aromatic residues.

Xie et al. in 2008 [38] experimentally studied the affinity of aromatic amino acids with a SWCNT and have found Trp to have the strongest affinity for both bundled nanotubes and individual nanotubes while Phe and Tyr have affinity only for the individual nanotubes. The binding strength among these three aromatic amino acids follow the order Trp>Tyr>Phe. Similarly, Rajesh et al. [39] have also observed the same trend in the strengths and found His to have the least binding strength among the aromatic amino acids in a DFT study. He and Zhou [40] in 2014 used MD simulations of a system containing all 20 standard amino acids and a SWCNT and identified Arg to also have a strong affinity along with the other aromatic amino acids. However, for a

graphene-GXG tripeptide system Camden et al. [41] have identified that Arg, Glu, Asp to have the highest binding enthalpy using MD simulations. Hughes and Walsh [42] have determined the free energy of adsorption of all 20 standard amino acids with graphene and found that Arg, Tyr, Trp and Gln to have the strongest and Ile to have the weakest binding free energies. In the case of a graphene oxide-peptide system, in an experimental study [43] a combination of π-π stacking and electrostatic interactions have dominated the driving forces as seen in the binding strength order of the amino acids Arg>His>Lys> Trp>Tyr>Phe. Recently [44] Arg, a positively charged residue was also reported to drive the interaction. Specifically, Wu et al. [45] have found that guanidinium ion (charged part) was found to have a major role in the interaction compared to norvaline (aliphatic part) part of the arginine residue.

The significance of water was pointed out in the interaction of proteins and CNMs, specifically in the DFT simulations where it is usually carried out in gas phase. Anversa and Piquini [46] have used tight binding DFT simulations, where they use configurations of water obtained from MD simulations in a DFT study. Among the amino acids studied, they found that the water has a significant role and in general the binding energies of non-polar residues is higher on a SWCNT. Lu et al. [47] have seen that water could impair the π-π stacking interaction. Shen et al. [48] studied the effects of charged SWCNT on insulin peptide adsorption and observed that, compared to neutral SWCNT, water is more ordered near the charged nanotube and effected in more compact adsorption of the peptide. Due to the presence of electrostatic interactions, the overall interaction energy was stronger in the case of a charged nanotube. Penna et al. [49] have

performed a statistical analysis on the interactions of peptides with graphene. They observed trends in the change in the driving forces from electrostatic to hydrophobic forces from the bulk to the surface by using data from a range of simulations of a peptide made of three domains (hydrophobic, hydrophilic and aromatic). Initial impetus to the adsorption was found to be dominated by the hydrophobic domain while the aromatic domain took control once the peptide is close to the surface compared to the hydrophilic domains, which did not have any significant role. Cooperative effects of both hydrophobic residues and hydrophilic residues were also observed in the simulations performed on proteins (albumin subdomain, fibronectin module) and graphite surface by Raffaini and Gannazoli [50], [51] in an implicit solvent. The adsorption behavior using the principles of thermodynamics was analyzed by Mijajolnic et al. [52]. The have decoupled the free energy into enthalpy and entropic components using MD simulations and found that the adsorption is favorable with respect to both the components but dominated by the former.

The role and stability of secondary structure content was also found in different studies, where in some of the cases α helix was found to be more stable and in some other cases beta sheets to be more stable [53], [54], [55], [56]. The dependency of the length of the α helix was shown in a different study [57], where contrary to the less stable shorter α helices, longer α helices underwent significant conformational changes. These studies show the importance of the sequence and the length of the peptide and in general the three dimensional structure of the protein in the interactions with solid CNMs.

Driving forces were also found to be dependent on the size and shape of the nanomaterial. Karunwi et al. [58] did not find a correlation with the type of amino acids on the adsorption behavior but saw differences in the interaction of the peptides with nanotube and graphene. In general, decrease in the curvature causes [59], [60] an increase in the adsorption for the studied proteins. Raffaini and Gannazoli [27] have also reported that compared to the outer convex surface of a nanotube, inner concave surface has a greater binding strength. As reported in section 2.2, defects can also have a role in influencing interaction between proteins and nanomaterials. However, when Walsh and Tomasio [61] used MD simulations to study oxidation defects on nanotubes, they found that there is no significant impact on the binding strength and the defects did not seem to serve as sources of attractive sites. In our recent combined experimental and simulations study, we found that the presence of intrinsic defects can have a major effect on the interaction energy and the corresponding conformational changes of the protein [62].

Table 2.2: Summary of the conclusions of an albumin-CNM system

System	Main Conclusions	Method
Albumin domains on a flat graphite surface.	Cooperative effects, both hydrophobic and hydrophilic interactions play a role, but the former interactions are stronger with the plane.	*in silico, (*MD Simulations) Implicit Solvent [55]

Albumin domains and SWCNT's with varying diameter	Non-polar residues (Val, Phe), polar and charged residues that contain the (-CH_2-) alkylene chain are near the SWCNT after interaction.	*in silico, (*MD Simulations) Explicit Solvent [59]
Bovine serum albumin on a flat graphite surface	Significant changes in the secondary structure content after adsorption.	*in silico, (*MD Simulations) Implicit Solvent [51]
Bovine serum albumin and a SWCNT	Among the four blood serum proteins studied, albumin has the least binding capacity and in all cases aromatic residues play a key role.	*in silico and in vitro* (MD Simulations, AFM, SDS-PAGE, CD and fluorescence spectroscopy [26]
Human serum albumin and a graphite surface	AFM images of individual albumin molecules and formation of network like and irregular structures at different pH is shown.	*in vitro* (AFM) [63]
Bovine serum albumin and MWNT with	Adsorption capacity of proteins on carbon based nanomaterials	*in silico and in vitro* (MD Simulations, Fluorescence

different radii, graphite surface	increases as the surface curvature decreases.	Spectroscopy) [60]

To summarize driving forces are not just dependent on hydrophobic and π- π stacking interactions for proteins-CNMs systems as reported in some of the recent studies and are also dependent on other factors depending on the protein and CNM.

2.4 Conclusions

Nanotoxicity, the study of toxic effects of nanomaterials has received wide attention over the past few years, due to the inevitable increase in chances of exposure of nanomaterials to the biological environment with the rapid rise in nanotechnology. The reason behind the toxic effects of nanomaterials is their extremely small dimensions, and their unknown impact on the biological molecules of similar dimensions such as proteins. In this chapter, we have summarized the recent reports on the factors and the governing forces that influence the interactions of biological environment with CNMs, which are the nanomaterials of our interest. The factors responsible could be divided into three categories, a) physicochemical properties of nanomaterials b) conditions of the biological environment and c) characteristics of the proteins. Physicochemical properties of the nanomaterials include the size, shape, stability in aqueous media, defects, impurities and surface modification while the biological conditions include solvent (water), temperature, pH, pressure and the presence of other biological molecules. Lastly, the type of amino

acids, sequence, secondary structure content, disulfide bonds, and three-dimensional structure comprise some of the characteristics of proteins that influence the protein-CNMs interactions. Compared to the numerous factors that influence the interactions of proteins with the biological environment, the governing forces are hydrophobic interactions, electrostatic interactions, and π-π stacking of aromatic side chains of proteins with CNMs. There is a disparity in what is commonly believed to be the most dominant driving forces. While some of the studies indicate hydrophobic interactions, π-π stacking to be dominant, some other studies indicate the significance of the other factors and a combination of factors mentioned above which include the role of water and the type of proteins. In such a highly complex system where different driving forces play key roles depending on the conditions and the type of the system, decoupling the factors influencing the protein-CNMs interactions is a challenging task.

CHAPTER THREE

ROLE OF SHAPE AND SURFACE CHEMSITRY IN THE INTERACTION BETWEEN ALBUMIN AND CARBON NANOMATERIALS

3.1 Introduction

Carbon nanomaterials (CNMs) are gaining attention lately due to their innumerable applications. The global market for carbon nanotubes alone is predicted to hit USD 3.42 billion by 2022 [64] and this growth indirectly implies increase in chances of their exposure to the biological environment. In the previous chapter we have seen that many physicochemical properties of nanomaterials play an important role in influencing the interactions between proteins and nanomaterials. Similarly, the biological conditions and the type of proteins also impact the interactions. In our study we focus only on two of the physicochemical properties of the nanomaterials a) shape and b) surface chemistry. We do it by considering the interaction between the most abundant blood serum protein, BSA [65] and CNMs comprising a single walled carbon nanotube (SWCNT), a graphene nanoribbon (GNR) and a graphene oxide nanoribbon (GONR). The width of GNR and GONR is the same as the circumference of the nanotube. SWCNT and GNR share the same surface chemistry (sp^2 hybridized) but differ in their shape while GNR and GONR share the same shape but differ in their surface chemistry. Albumin, though found to have the least affinity to the CNMs in some of the recent studies [26], due to its high concentration in the blood serum has the highest chances in making a contact with CNMs, especially when intentionally exposed in applications where CNMs are used as

drug carriers. Thus studying such a system comprising albumin and CNMs would also provide insights into the significance of shape and surface chemistry into the toxic affects of nanomaterials. We employ all atom molecular dynamics simulations using state of the art techniques to meet the computational demands in probing these interactions. To the best of our knowledge, this is the first study of such a system comprising the entire albumin chain, explicit water model and the three CNMs.

3.2 Methods

We employ molecular dynamics (MD) simulations to explore the interactions between proteins and CNMs at a molecular level. This computational method along with the principles of statistical mechanics is an excellent tool for studying such systems which is otherwise not feasible using available experimental methods. It provides the trajectory of each and every particle that constitutes the system allowing one to study the parameters influencing the behavior of the system. This method obeys classical "Newton's law of motion" and uses it to predict the future of the system. The governing forces between the particles in this scale of simulations are divided into a) bonded interactions and b) non-bonded interactions. The parameters involved in describing these interactions distinguish each particle or a group of particles and are obtained either from experiments or from ab initio calculations. The validity of these parameters is tested for various systems by testing the results of the MD simulations with experiments and has been generalized by different research groups. These sets of parameters that describe the interactions are called as force fields. The only initial requirement for these simulations is

a good starting configuration of the system and an appropriate force field. Below we first present a general algorithm of MD simulations [66], [67] and then describe the interactions that govern the forces between the particles comprising the system.

MOLECULAR DYNAMICS ALGORITHM

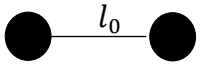
- Input Initial Conditions (starting configuration)
  - Positions and Velocities of all particles (i) in the system
    $$r_{i,}\, v_i \text{ at } t = 0$$
  - Force field describing interactions between all particles as a function of positions $\left(V(r_i)\right)$
- Compute Force
  $$F_i = -\frac{\partial V(r_i)}{\partial r_i}$$
  where
  $$V(r_i) = V_{bonded}(r_i) + V_{non-bonded}(r_i)$$
- Update Configuration
  - "Newton's law of motion"
    $$F_i = m_i a_i = m_i \frac{d^2 r_i}{dt^2}$$
  - Numerical solvers to solve for positions and velocities
    $$\frac{dr_i}{dt} = v_i;\ \frac{dv_i}{dt} = \frac{F_i}{m_i}$$
- Write Output Configuration (r_i, v_i)

Repeat required number of times

FORCE FIELD

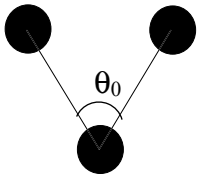
- Bonded interactions
  - Bond stretching interactions

$$V_{bond} = \sum_{bonds} k_b(l - l_0)$$



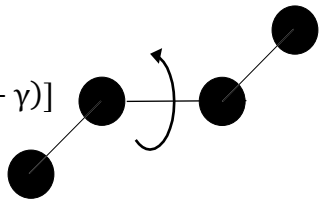
  - Bond angle bending interactions

$$V_{angle} = \sum_{angles} k_\theta(\theta - \theta_0)$$



  - Dihedral angle interactions

$$V_{dihedral} = \sum_{dihedrals} \sum_{n} \frac{1}{2} V_n[1 + (\cos(n\omega - \gamma)]$$



- Non bonded interactions
  - Van der Waals (vdw) forces

$$V_{vdw} = \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left\{ \epsilon_{ij} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^{6} \right] \right\}$$

  - Electrostatic forces

$$V_{electrostatic} = \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left\{ \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}$$

Finally,

$$V = V_{bond} + V_{angle+} V_{dihedral} + V_{vdw} + V_{electrostatic}$$

Where:

i, j represents particles i and j. N is the total number of particles.

r_{ij} is distance between i and j, ϵ and r_0 represent vdw parameters.

q is the partial charge on a particle and l_o, θ_0 represents equilibrium bond length and angles.

n represents the number of terms in the Fourier series. $\omega \; and \; \gamma$ represents the dihedral angles and phase offset.

f_{ij} is a scaling factor that scales down all 1-4 non-bonded interactions.

All other terms represent the parameters that are defined by the force field depending on the type of particle.

Note:

1. For bonded atoms, non-bonded interactions are usually calculated between atoms that are three or more bonds apart.
2. $\varepsilon_{ij} \; and \; r_{ij}$ is calculated based on standard set of combination rules (arithmetic mean or geometric mean).
3. Amber uses the above form of force field [].

3.3 Simulation Setup

As described in the above section, we need a good starting configuration for setting up a simulation. This starting configuration is usually determined from experiments or modeled based on the available experimental properties of the molecule. In general, the protein data base bank (RCSB PDB) [68] maintains the three dimensional

structure of all the proteins as and when they are determined. The starting configuration of bovine serum albumin is obtained from this database, where its three dimensional crystalline structure is reported (4F5S.pdb) [65]. The nanotube is modeled with (16,0) zigzag SWCNT configuration and the width of GNR and GONR is equal to the circumference of the SWCNT. The starting configurations of GONR is obtained using the methodology described in DeFever et al. [69]. In all our simulations oxidation of GONR is fixed to 30%. Figure 3.1 shows the starting conformations of each of these CNMs and BSA.

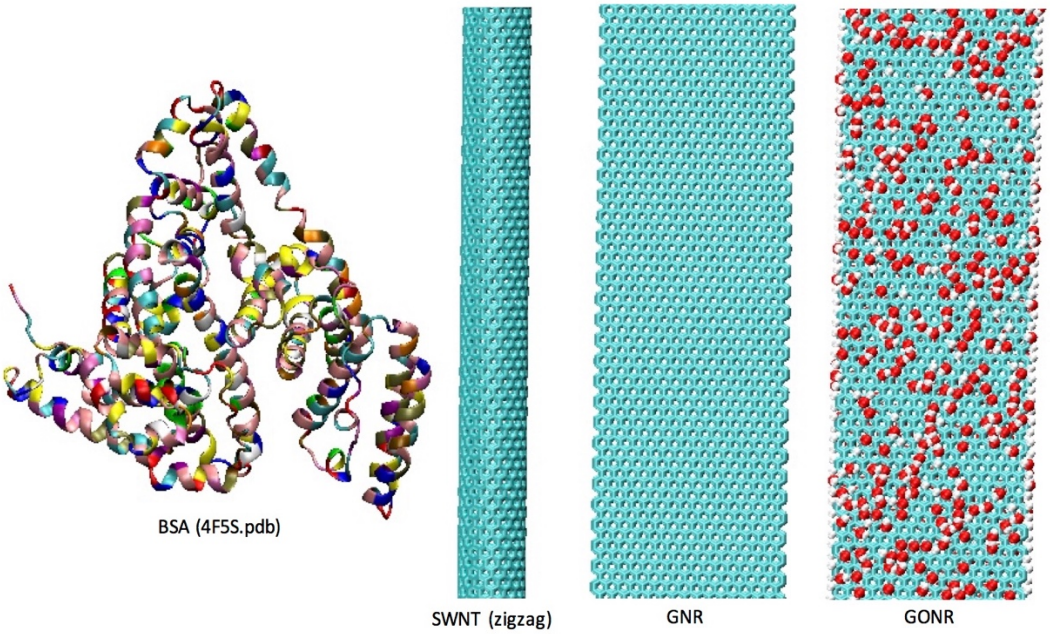


Figure 3.1: Starting structures of BSA (represented by new cartoon), SWCNT, GNR and GONR (represented by licorice model) respectively. Oxygen atoms and hydrogen atoms on GONR are represented by vdw model.

We performed large-scale MD simulations of BSA-CNM in explicit water using TIP3P water model [70]. The Amber ff99SB-ILDN [71] force field was used to describe the protein and SWCNT/GNR while GONR was described using the OPLS [66] force field because it eliminates the necessity of determining the partial charges on the GONR atoms using ab initio calculations. Each of the simulation systems comprises of the CNM atoms, 9220 protein atoms and ~70,000 water molecules and ions that were added to neutralize the system. Fig 3.2 shows the set up of such a system.

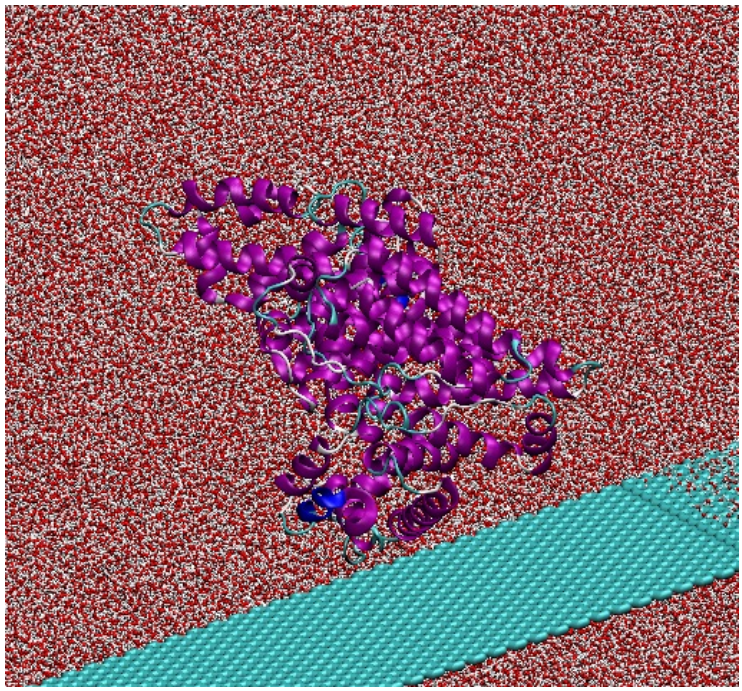


Figure 3.2: Schematic showing the initial set up of a BSA-GNR system with TIP3P water molecules. Some of the water molecules are not shown for clarity.

In order to enhance the sampling of protein adsorption in the MD simulations we performed 10 simulations for each system with 10 different protein orientations in the starting configuration of the simulation. These starting orientations are obtained by choosing one reference orientation and rotating it uniformly by 30^0, 60^0, 90^0 about each axis as illustrated in Fig 3.3.

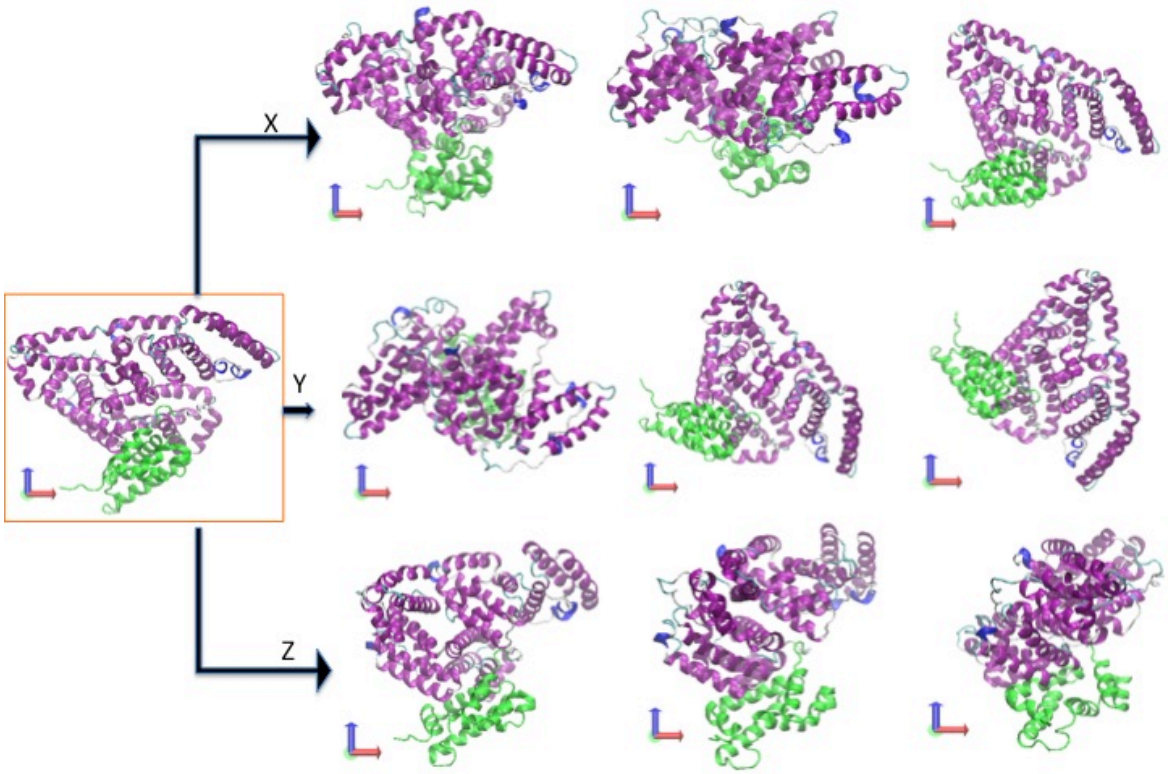


Figure 3.3: Initial orientation of BSA in each of the BSA-CNM system. Same domain is highlighted in green for all orientations to provide a perspective in the differences between each orientation.

In this way, we could sample the interactions of various regions of BSA with the CNMs within accessible time scales. We note that, this does not capture all the possible regions of the protein or the regions that have an affinity to the CNM. Initially, the protein was placed such that no two heavy atoms of the protein and the CNM were less than 8 Å. The velocity-rescaling thermostat [72] and Berendsen barostat [73] were used to maintain the temperature at 300 K and pressure at 1 bar, respectively. The long-range electrostatic interactions were calculated using PME [74] as implemented in GROMACS v 5.0.2 [75] . The positions of all the CNMs are fixed and the bonds involving hydrogen atoms were constrained using LINCS algorithm [76] and time step used was 2 fs. In total, we have performed 100 ns simulations resulting in a net total of 1 μs simulation per BSA-CNM system. All simulations were run using GROMACS v5.0.2 with GPUs (Appendix A) and configurations are stored every 20 ps for each of these systems for further analysis.

3.4 Results and Discussion

3.4.1 Interaction energy

We quantify the initial adsorption behavior of BSA by first determining its interaction energy with CNMs. The interaction energy is the combination of Van der Waals and electrostatic forces between BSA and the CNMs. It excludes all other self-interactions and interaction of BSA, CNMs with the solvent providing an estimate of the strength of the interaction between BSA and CNMs alone. Figure 3.4 shows the changes in interaction energy of BSA with each of the CNMs during the course of the simulation

for three of the starting configurations. For each of these configurations, the interaction energy decreases in a step-wise manner. The overall drop in this interaction energy is highest for the BSA-GNR system compared to the other two systems. This implies that the interaction of BSA is stronger with GNR compared to SWCNT and GONR.

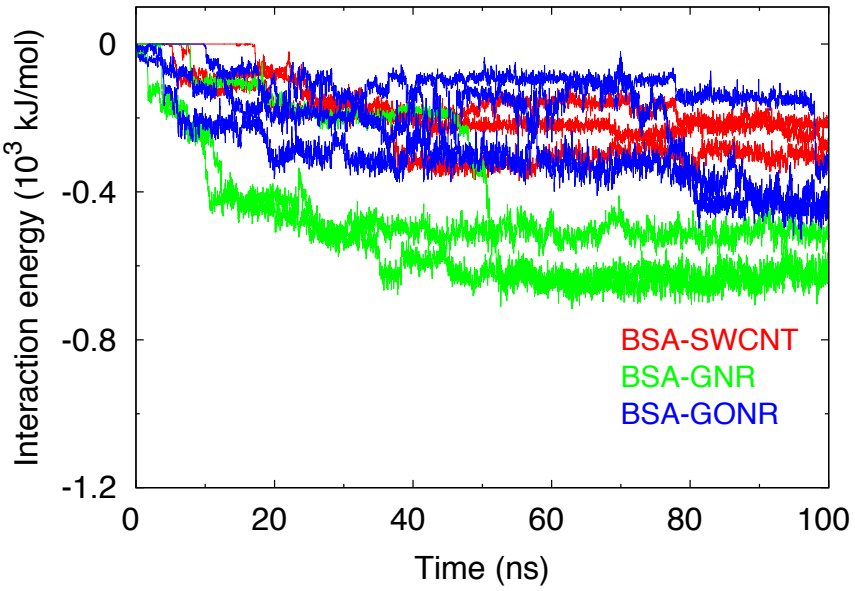


Figure 3.4: Change in interaction energy of BSA with different CNMs during the course of the simulation. Each line corresponds to a simulation with different starting configuration. Appendix B includes interaction energy plots for all the starting orientations.

A similar trend in the interaction energy is observed for varying curvature of CNMs with identical surface chemistry was also observed in other studies [59], [55], [60]. In particular, Gu et al. [60] considered an armchair SWCNT contrary to our zigzag conformation, have also observed such a trend. The weaker interaction of GONR, which

only differs in the surface chemistry when compared to GNR suggest that hydrophobic surfaces interact strongly with BSA. In order to probe into the regions of steep drops, we have calculated the number of heavy atoms of BSA within 5 Å of the CNMs. We have observed a strong correlation between the interaction energy and the number of heavy atoms near the CNM.

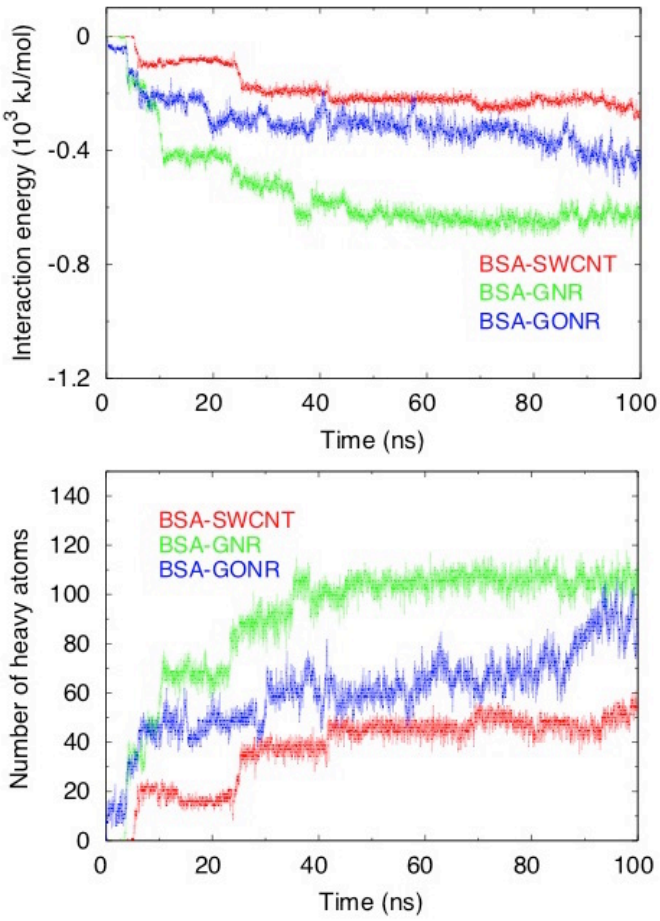


Figure 3.5: Interaction energy and the corresponding change in the number of heavy atom of BSA near CNMs.

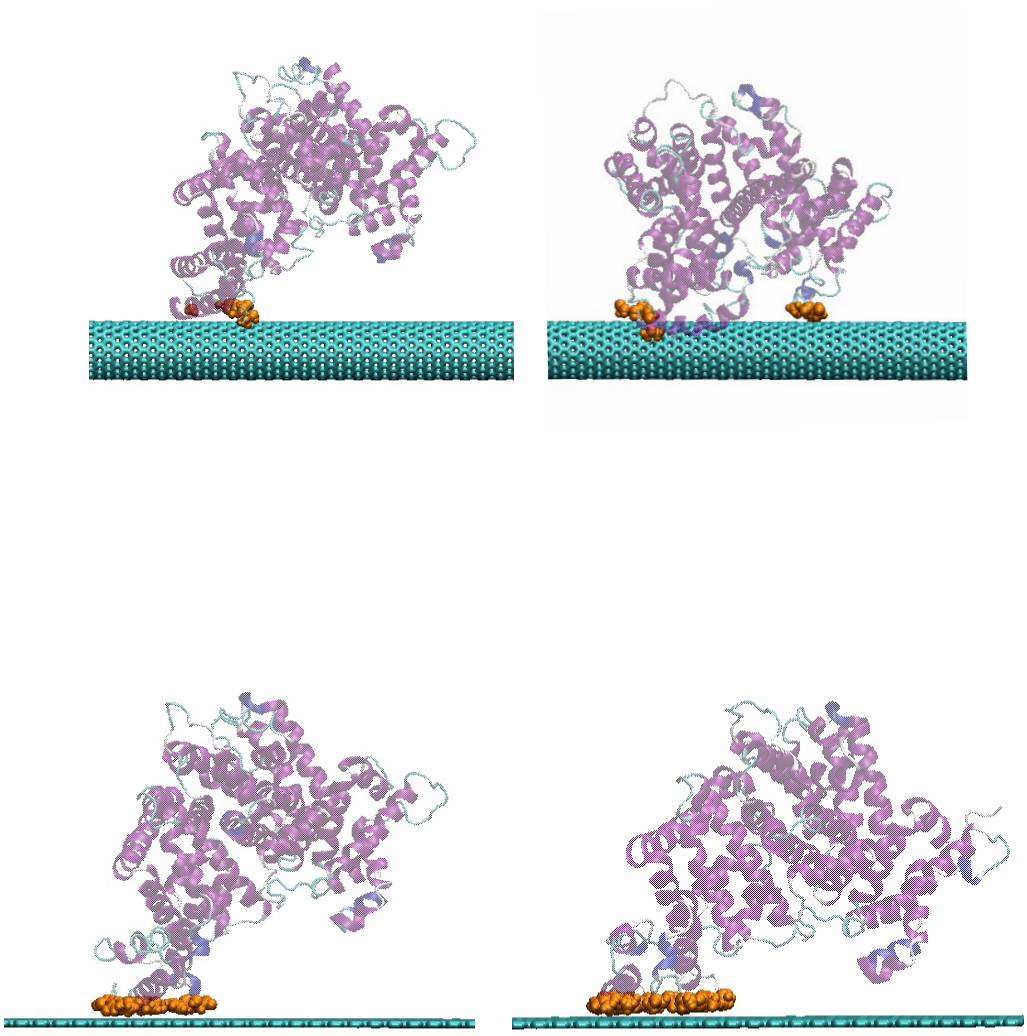


Figure 3.6: Snapshots of the initial adsorption of BSA with SWCNT and GNR. All the snapshots on the left correspond to t=10 ns and all the snapshots on the right correspond to t=35ns.

In Fig 3.5, we show the interaction energy for one of the starting orientation and the corresponding numbers of atoms near the CNM. The strong correlation implies the abrupt drop in the interaction energy is an abrupt increase in the number of atoms of BSA near the CNM. An increase in number of contacts of BSA is possible when there is a change in the conformation of BSA. This change could be either in the vicinity of the CNM or when some other segment of BSA comes closer to the CNM. These possibilities are shown in Fig 3.7, which also corresponds to the drops in the interaction energy for GNR and SWCNT. These trends in the interaction energy is also consistent with other recent reports on the interaction of albumin with SWCNT, MWNT and graphene.[51],[55],[26]. However to the best of our knowledge, this is the first study that compares the interaction of BSA with CNMs of varying shape and surface chemistry.

3.4.2 Secondary structure content of BSA

Bovine serum albumin is predominantly α helical and contains three domains with similar conformations [65] in its crystalline structure. Each of these domains contains two subdomains A and B and contains six helices each as highlighted by different colors in Fig 3.7 a. We have performed a simulation of BSA in water and found that this structure is also stable prior to running the simulations with CNMs. Fig 3.7 b includes the structure of BSA in water after 10 ns. In the previous section, we concluded by suggesting that during the initial adsorption process, BSA undergoes conformational changes in the presence of CNMs. To quantify the changes in secondary structure content due to these conformational changes, we calculated the secondary structure content during the course

Figure 3.7: a) Illustration of subdomains of the BSA (represented by new cartoon model and each color represents a sub domain) and b) Final structure of the BSA after 10 ns simulation in water (represented by new cartoon and magenta represents α helices, blue represents 3-10 helices, white represents coils and cyan represents turns)

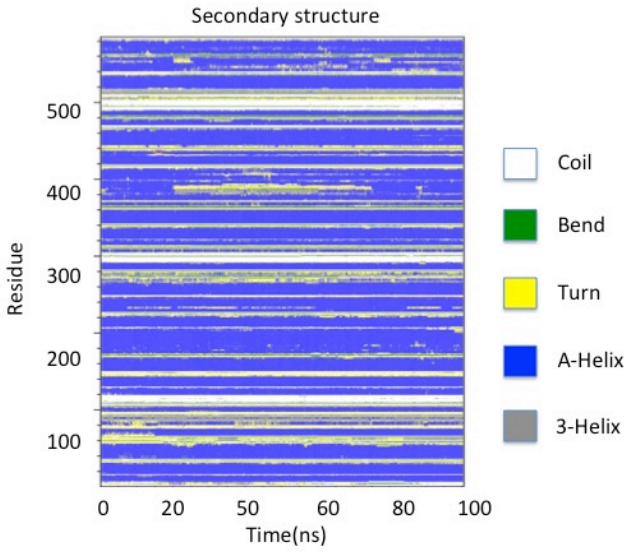


Figure 3.8: Secondary structural changes of BSA during the course of simulation with GNR.

of the simulation. Fig 3.11 shows the changes in the secondary structure content of BSA-GNR system obtained by using dictionary of protein secondary structure (dssp) [77]. As observed, the changes in the secondary structure content are not significant showing the stability of α helices in the initial adsorption. This is also true for other BSA-CNM systems for all starting orientations during the course of our simulations. This also suggests that all the conformational changes in BSA are only changes in the relative rearrangement of the domains of BSA. These changes in secondary structure content are completely in contrast with the recent experimental reports [55], [26] and this is due to the differences in the time scales involved but also due to the involvement of other factors such as presence of nanotube bundles, concentration of proteins, and biological environment which can be controlled with precision in simulations.

3.4.3 Role of water in the interaction of BSA with CNMs

In Fig 3.12, we report the normalized number of water molecules (N_{hyd}) in the hydration shell of radius 5 Å in the BSA-CNM complex during the course of the simulation for one of the starting orientation reported in Fig 3.6. As expected for a hydrophilic surface like GONR, there are more water molecules in the vicinity of the surface compared to hydrophobic surfaces SWCNT/GNR. However, the positive correlation in the drop with the interaction energy in Fig 3.6 for all CNMs implies the overall decrease in the number of water molecules, as BSA adsorbs on CNM. We also calculated the contact surface area defined as the difference between solvent accessible

surface area of BSA alone and BSA-CNM complex, reported in Fig 3.12. Here again, we observed a strong correlation between the contact surface area and interaction energy,

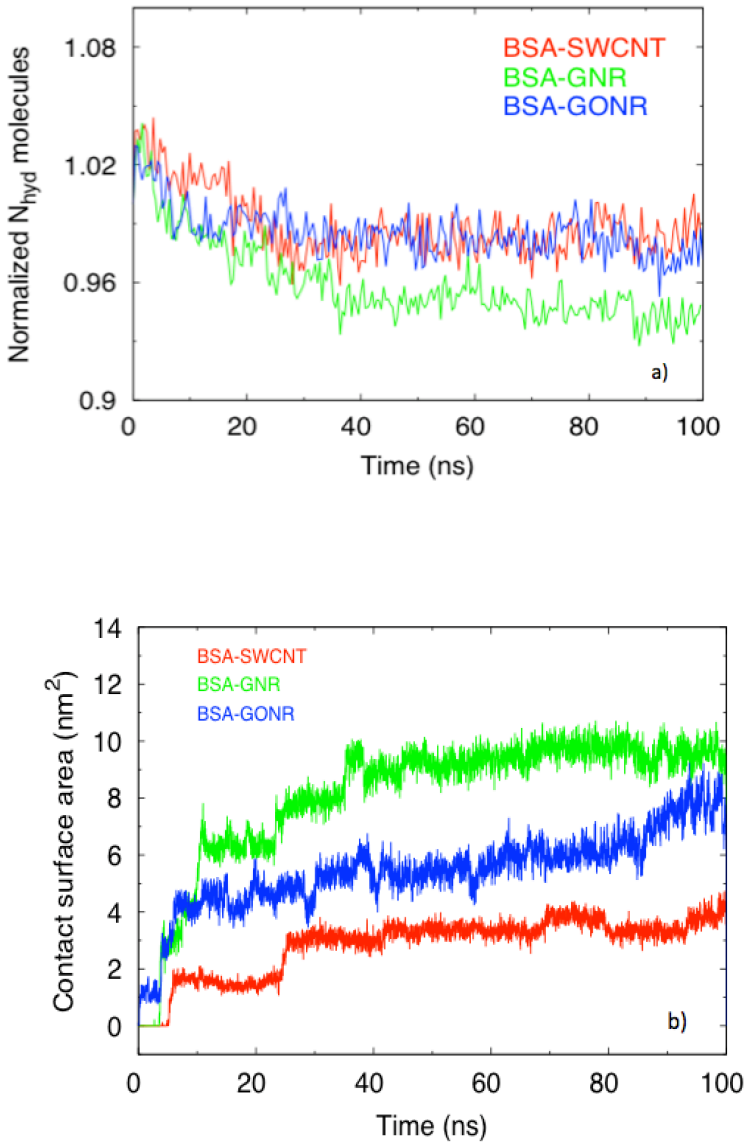


Figure 3.9: a) Change in the normalized number of water molecules in a shell of radius 0.5 nm around BSA-CNM complex and b) Change in contact surface area of BSA with CNM during the course of the simulation.

which indicates the absence of water molecules between the two molecules during the interaction. These strong correlations suggest the role of water during the interaction of BSA with CNMs.

3.4.4. Amino acids near the CNMs

In order to quantify the driving forces, we determined the normalized time averaged number of amino acids near the CNMs in the last 20 ns of BSA-CNM interactions and reported in Fig 3.13. Here we presented for the same three starting orientations reported in the interaction plot in Fig 3.4 and reported the data collected for others in Appendix B. An amino acid is considered to be near the CNMs when 80% of its heavy atoms are within 5 Å of the CNM. The amino acids are categorized based on their type of side chains. As observed in Fig 3.13, when different regions of BSA are exposed to the CNM (starting orientations), different amino acids are found near the CNMs. For the orientations shown here and others reported in Appendix B, we do not observe the dominance of any one or two amino acids of a particular type of side chain when BSA interacts with CNMs. However, the presence of glycine for all CNMs and starting orientations suggest the higher interaction of coils and loops in BSA with the CNMs than the predominant α helices. In the case of BSA-GONR system, for most of the orientations there are no amino acids with aromatic side chains suggest the absence of π- π stacking interactions. This is expected due to lower availability of aromatic rings in GONR compared to GNR and SWCNT. The lack in the dominance of aromatic side chains interacting with the CNMs suggest that π- π stacking interactions may not be the only

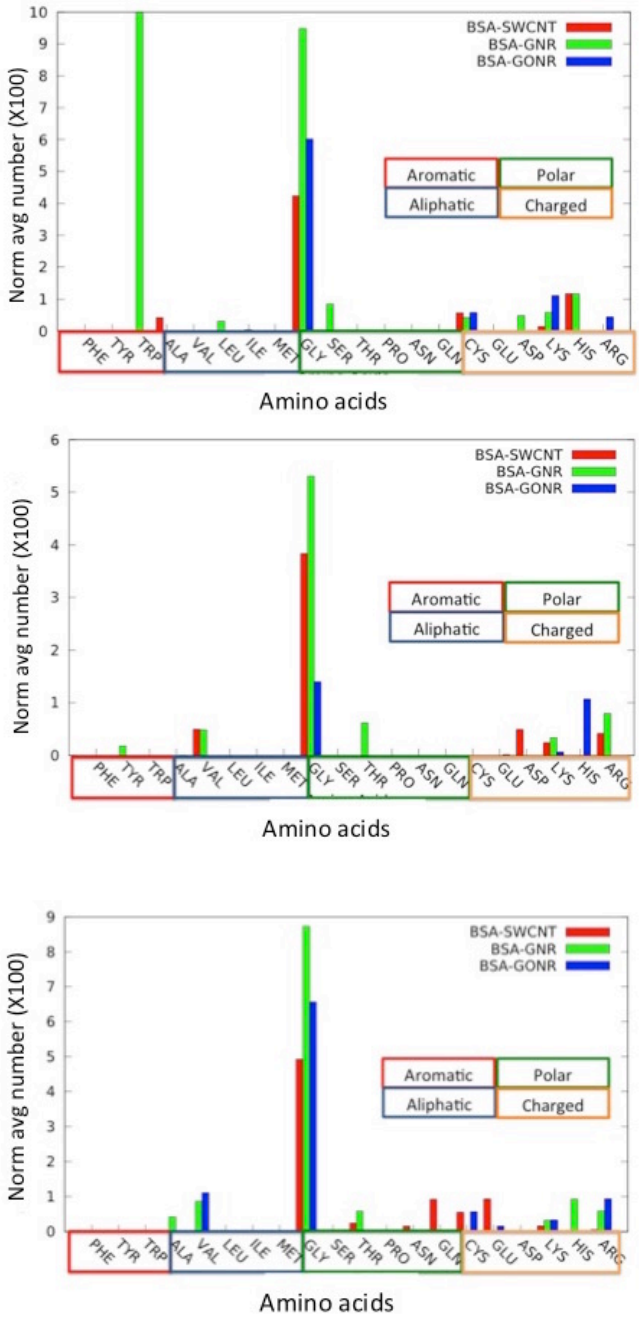


Figure 3.10: Normalized time average of the number of amino acids during the last 20 ns for each of the BSA-CNM system.

dominant driving forces in the interaction of BSA with CNMs as reported in some of the recent studies [26], [60].

3.4 Conclusions

In this chapter, we studied the effects of shape and surface chemistry of CNMs in its interaction with BSA. We have observed that, all the CNMs interact with varying interaction strength and follow the order GNR>SWCNT~GONR. The lower interaction strength of BSA-SWCNT compared to BSA-GNR suggests that the interaction of BSA with CNM increases with decrease in curvature. The secondary structure calculations show that the conformational changes during BSA-CNM interactions does not result in significant changes in the predominant α-helical content of BSA during our course of simulations. We observed positive correlation of the number of water molecules in the hydration shell of BSA-CNM complex with the interaction energy suggesting the role of water in the initial adsorption process. In all the BSA-CNM interactions we did not observe the dominance of any one or two amino acids of a particular type of side chain during BSA-CNM interactions. This indicates the presence of other factors and not just the type of residue in the interaction of BSA with CNMs.

CHAPTER FOUR

ROLE OF A NEIGHBORING RESIDUE OF AN AMINO ACID CONTAINING AROMATIC SIDE CHAIN IN THE INTERACTION OF PROTEINS WITH CARBON NANOMATERIALS

4.1 MOTIVATION

In the previous chapter during the interaction of BSA with CNMs, there was no dominance of any one or two amino acids of a particular type of side chains. In particular there are no aromatic groups near the CNMs for many of the starting configurations of the BSA. However, recent studies as reported in chapter 2 of this thesis suggest varying types of driving forces between proteins and CNMs. These include the dominance of aromatic side chains, arginine, water and a combination of polar and non-polar side chains. In order to quantify these factors we designed a set of simulations that would not only show the role of an aromatic side chain/ arginine in its interaction with proteins but would also provide us the role of water and the role of neighboring residues of an aromatic or arginine residues. These simulations include a set of tripeptide-CNM systems. We probe these factors by determining the free energy of interactions of proteins with CNMs using umbrella-sampling method. In this chapter, we report only the results of the tripeptide system where the central residue is phenylalanine and the other systems are a part of the proposed future work.

4.2 Methods And Simulation Setup

We employ advanced sampling method, umbrella sampling to calculate the free energy of interactions of a set of tripeptide-CNM systems. This method enables us to obtain the free energy landscape by restraining the system with a harmonic potential along an order parameter. Below, we provide the general algorithm of umbrella sampling method and Appendix B provides the details of the procedure to run these simulations using GROMACS with an example.

4.2.1 Umbrella sampling

Umbrella sampling is an advanced sampling method that enables us to obtain the free energy landscape of the system and can be readily applied to molecular dynamics (MD) simulations. Torrie and Valleau first suggested this method in 1977 [78]. We explain this method by providing an example based on our system, protein and a CNM. Consider an arbitrary surface and solute that are interacting as shown in Fig 4.1. In the previous chapter, when we had such a system where the solute is bovine serum albumin (BSA) and the surface is a CNM, we had seen that the simulation is quite expensive and the sampling was limited. Fig 4.2 shows an example of a free energy landscape of a protein. As seen, the landscape is quite rugged and crossing the energy barriers to obtain different possible conformations is quite difficult. A traditional MD simulation would take a long time and a large number of computational resources are required to access configurations at all points on the free energy landscape. In order to overcome this problem umbrella sampling method obtains the conformational spaces in a discretized

manner. It first samples all the conformations at different values order parameters and then combines them to obtain the overall conformational landscape of the system.

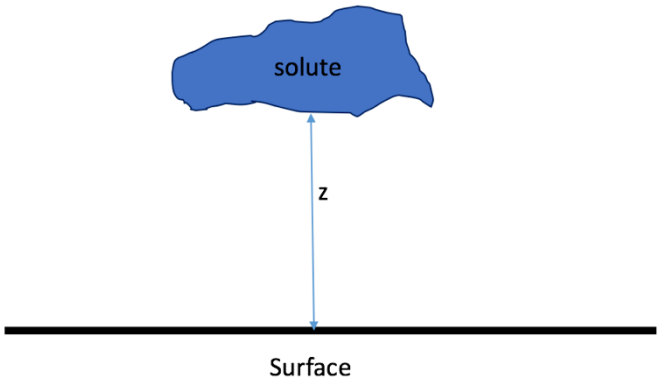


Figure 4.1: Illustration of umbrella sampling simulation where the z distance between the solute and the solvent is the order parameter.

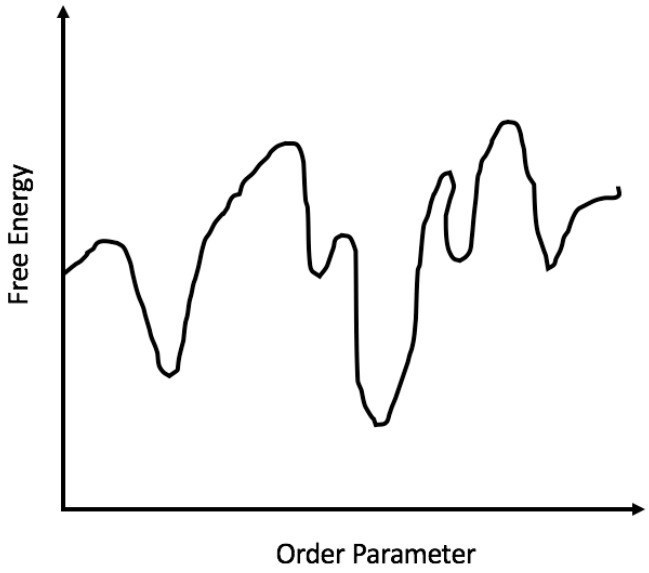


Figure 4.2: Illustration of a rugged free energy change with an order parameter.

In Fig 4.1, the order parameter is the z distance between the center of mass of the molecule and the surface. In this way, when the molecule is restrained to a particular distance, all the possible conformations of the molecule and the surrounding solvent molecules are obtained. This process is repeated for different values of order parameter (z distance in the example) in order to measure the distribution of the order parameter the molecule accesses during the process. An example of such a distribution is shown in Fig 4.3. The unbiased distribution is obtained by removing the added biased potential and the final free energy of interaction is obtained by using a weighted histogram analysis (WHAM) [79] method.

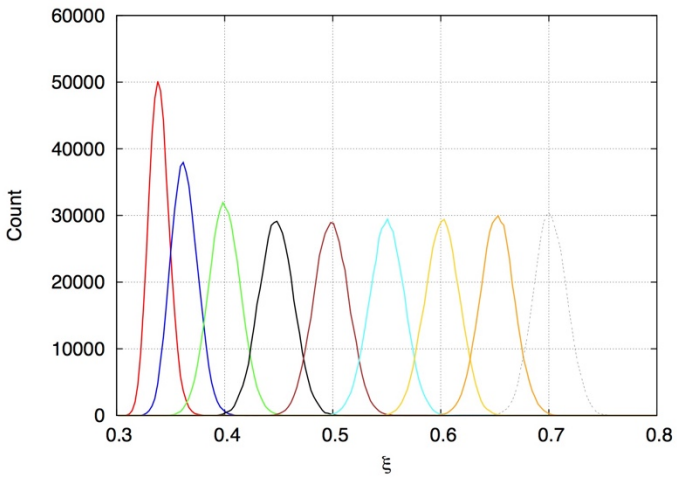


Figure 4.3: Distribution of order parameter obtained after umbrella sampling simulation.

Note:

- Umbrella sampling method can be applied to arbitrary reaction coordinate and the example illustrates the order parameter to be the z coordinate. Indirect sampling

method (INDUS) uses the number of water molecules as the order parameter to study water density fluctuations [80].

- It is quite important for the windows or the range of order parameters chosen in order to obtain an accurate potential of mean force

4.2.2 Simulation setup

We use the above procedure to calculate the free energy of interaction for a tripeptide-CNM system. The tripeptides are designed such that the central residue is an amino acid containing aromatic side chain or arginine. The aromatic groups of interest are phenylalanine and tryptophan based on their strength of interaction among aromatic residue shown in various studies. One end of the tripeptide is fixed with glycine, as it does not have a side chain. The other end is the neighboring residue of interest, which have varying size of side chains and include both polar and aliphatic residues.

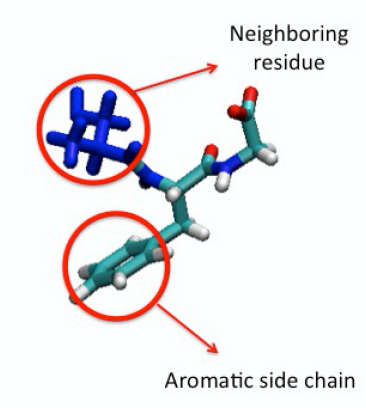


Figure 4.4: Schematic illustrating the tripeptide system (represented .by licorice model)

In other words, the tripeptides are of the form X-Phe-Gly, X-Trp-Gly, X-Arg-Gly and X-Phe-X where X is (Val, Leu, Gly, Thr, Ser, Arg) as shown in Fig 4.6. We initially study only GNR as the CNM and other CNMs are a part of the proposed future study.

For the umbrella sampling simulations, we choose the reaction coordinate (ξ) to be the z_0 distance between the center of mass of the aromatic side chain of the central residue and the CNM as shown in Fig 4.5. The reaction coordinate is varied from 0.4 nm to 2 nm with a window spacing of 0.1 nm. Each window is run for 25 ns and repeated 5 times at a temperature of 300K and pressure of 1 bar. The biasing potential in each window and for each tripeptide is 3000 kJ/mol/nm^2 and the distribution of reaction coordinates is stored every 1 ps for obtaining PMF. All the simulations are performed using GROMACS v 5.0.2 and each of the tripeptide system is solvated by 1260 TIP3P model water molecules.

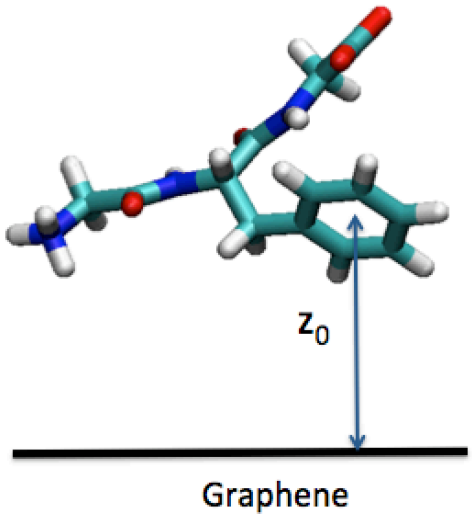


Figure 4.4: Initial set up of the tripeptide-graphene system

4.3 Results And Discussion

In the following section, we report only the potential of mean force obtained for the X-Phe-Gly system and include the other system in a later work. Fig 4.6 shows the umbrella windows obtained for one of the system indicating the windows are well overlapped and the PMF comparing all the systems is reported in Fig 4.7. Error bars are calculated by measuring standard deviation from the repeated five runs of each system. As seen, the PMFs for all the systems where X is Val, Thr, Ser, Leu and Gly are similar. This indicates the neighboring residues (Val, Thr, Ser, Leu) of Phe do not have a significant influence effect on the overall free energy of interaction of tripeptides with graphene. However, barriers appear during the interactions for the larger and polar side chains near 0.6-1.2 nm. We hypothesize this is due to conformational changes that the larger side chains need to make as the tripeptide associates with the CNMs.

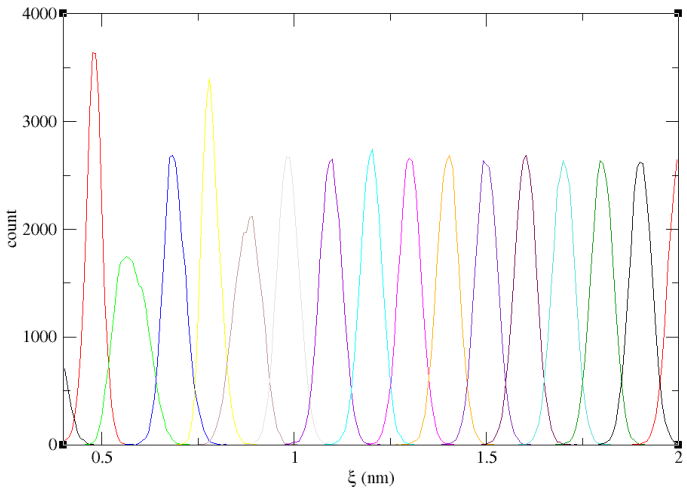


Figure 4.5: Distribution of order parameter after umbrella sampling simulation

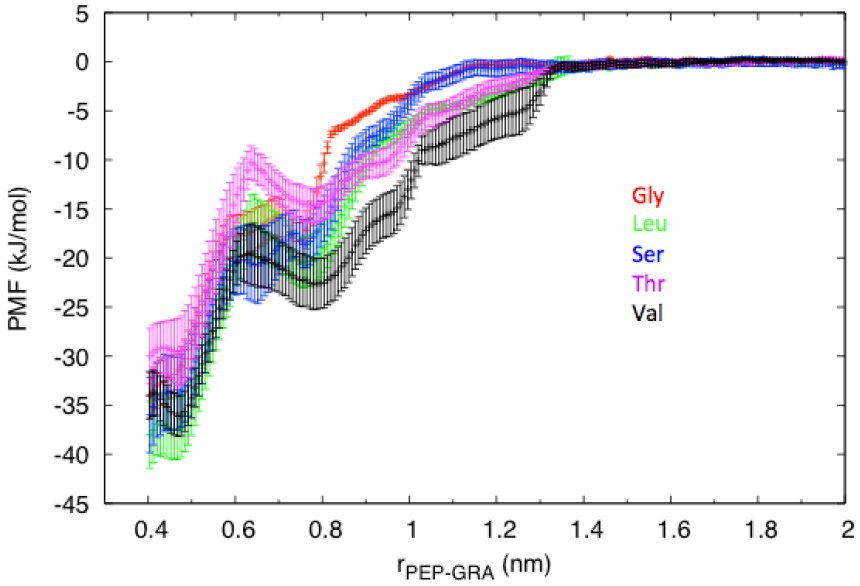


Figure 4.6: PMF of the X-Phe-Gly system with error bars.

4.4 Summary

In this chapter, we have presented a set of simulations that would enable us to determine the effect of neighboring residue of an amino acid containing aromatic side chain and arginine in its interaction with a CNM. In the fist section, we presented the methodology used in running these simulations and the details of the simulation setup. We presented the results for the X-Phe-Gly tripeptide where X is (Val, Thr, Ser, Leu and Gly) and found that the free energy of interaction is similar in all the cases. However, barriers appear during the interactions for the larger and polar side chains. We hypothesize this is due to conformational changes that the larger side chains need to make as the tripeptide associates with the CNMs. A part of the proposed future work includes exploring the role of water by decoupling free energy into enthalpy and entropic

components and also calculating the PMFs for the other aromatic groups, arginine and other CNMs. Recently, Arginine was shown to have an equal role in the interaction of proteins with CNMs and determining the free energy of the tripeptides Arg-Phe-Gly would enable us to find its strength compared to an aromatic group.

CHAPTER FIVE

SUMMARY AND PROPOSED FUTURE WORK

5.1 Summary

Engineered nanomaterials have wide array of applications, but there is growing evidence on their influence on the conformations of proteins. In this thesis, where we focus on carbon nanomaterials (CNMs), we first enlisted some of the factors that influence this behavior from recent experimental studies and computer simulations. Broadly the factors can be categorized into the physicochemical properties of the nanomaterials, the biological environment and the characteristics of the protein. The physicochemical properties include the size, shape, surface chemistry, defects and impurities while some of the characteristics of the proteins include its three dimensional structure, sequence, secondary structure content and stability.

We then provided the dominant interactions between proteins and CNMs and found that there is a disparity among the conclusions, where some studies suggest a combined role of hydrophobic and hydrophilic residues and others suggest dominance of a basic residue, arginine. Some of the recent studies on the simulations of CNMs and albumin, a dominant blood serum protein, have indicated the dominance of hydrophobic and π-π stacking interactions. However, in our simulations we have not seen such a dominance, but we found the interactions between BSA and CNMs to follow the order GNR>SWCNT~GONR. We also did not observe any significant changes in the secondary structure content of BSA suggesting α helices are stable during our course of

simulations. As a part of further analysis we would like to look into the local conformational changes by calculating the Ramachandran angles, which would show the changes in the orientations (if any) during the initial adsorption.

As the number of studies suggesting the dominance of π- π stacking interactions is increasing, we have designed a set of simulations to study the influence of the of aromatic residues over its neighboring residues. This set of simulations, which uses advanced sampling methods would enable us to determine the adsorbed free energy and provide us the most dominant aromatic group. These simulations could also allow us to study the influence of basic residue arginine. Our initial results suggest that the tripeptide systems where the central residue is Phe and the neighboring residues are (Val, Thr, Ser, Leu) have similar free energy of interaction. As a part of the proposed future work, we would also like to run the simulations at different temperatures to determine the enthalpy and entropy contributions to the free energy of interaction.

5.2 Proposed Future Work

The comparison of computer simulations with experiments to a full extent may not be possible all the time. This is due to presence of various other uncontrollable factors that effect experiments. However, computer simulations do provide important insights into the governing forces. In chapter 2 of this thesis, which includes a summary of the recent studies, we found varying factors that govern the interactions between proteins and CNMs. However, some of the possible factors that effect these interactions depending on the system are discovered generating a pool of data. Here, we propose a method, which

enables us to predict the governing forces and also other possible factors given the components and conditions of the system using this data. This method involves developing a mathematical model that involves the governing forces, the factors influencing the system and the components of the system.

The first step of this method is to find a set of valid data that could be used to build the model. One way of doing this is by determining all the computer simulations and the experimental results of similar systems from the literature. For example, first enlist all the types of nanotubes used in both simulations and experiments and then categorize those simulations, which have used similar nanotubes as experiments and have similar conclusions. In the case of varying conclusions, include the list in pool B and in the case of similar conclusions, include in pool A. The second step is to focus on pool A and sort the nanotubes based on their physicochemical properties and list them in column A of pool A, the proteins used in column B of pool A and finally the dominant interactions in pool C. We want to make a note that, the method is described to use the data available in the literature and the data that will be generated by us in future.

The third step is to map column A with column C for a same value in column B. As the conclusions in some of the studies for similar systems can be different, we might have different possible mappings. For example, mapping single walled nanotube with BSA and its dominant interactions. Once the process of mapping is complete, then our goal is to build a mathematical model, which would predict column C given column A and column B using this mapped data. As mentioned earlier, we are making an assumption on the availability of the data at this stage. If the model is able to predict the

result accurately, the latter step is to find methods that could enable us to create a model which can determine column C give column B and a value which is a combination of one or more values of column A. For example, assume one of the columns A is the diameter of the nanotube and the other value is defects of the nanotube. Now, given a combination of diameter and defects of the nanotube, we want to create a model that can predict the dominant factors (column C) given the protein (column B). In this way, theory can be combined to approach experiments as more and more combinations of column A would result in a structure that is closer to reality. This is otherwise difficult to study using simulations alone. Following tables illustrates the model development,

Table 5.1 A sample table illustrating the method.

Column A	Column B	Column C
A1	B1	C1
A2	B2	C2
A3	B3	C3
A4	B4	C5

Assume in table 5.1, Column A represents the set of physicochemical properties of nanotubes, Column B represents a protein and Column C represents the possible dominant factors influencing interactions of proteins and nanotubes. It is possible that A1, B1 combination can be mapped to a single value or multiple values in Column C. It is also possible the results from various studies to be same or different. In the case of

similar results, the combination would result in a single value and in the case of different results we would have different set of data that has to be used in the model development. At this stage, with sufficient set of data we think applying statistics might help obtain a relationship model. However, in the case of in sufficient data other mathematical models have to be explored. Once the model is able to predict the relationship with certain accuracy, a general algorithm should be developed in order to obtain the relationship when a combination of A1, A2 and B1 is provided in order to predict its relationship to column C.

As the data increases, the model gets better and might predict the type of interactions between a nanomaterial and a particular protein, which would help in developing safer and efficient nanomaterials. This methodology can be further extended to other nanomaterials and the model can be generalized for all set of nanomaterials and proteins. We note that the above-illustrated method may not be an easy task and the availability of data might be a hindrance. However, this method enables to organize the current area of research on nanotoxicology. It also assists researchers to perform studies in a systematic way compared to the current exploration of interactions where the conclusions are varying due to possibilities of different factors. Similarly, in order to explore all the factors, in future a mapping relationship can be developed which maps the system, factors and the governing forces.

APPENDICES

Appendix A

GROMACS WITH GPUS

In order to run large-scale simulations in a reasonable amount of time, we have used the available computational hardware (Tesla k20 model) in Clemson universities high performance computing (HPC) resources to increase the computational speed [81]. GROMACS is an open source MD simulation package and is of one the fastest parallel MD codes to run across multiple CPUs and GPUs [82]. However, at the time of running our all atom simulations mentioned in chapter 2 of this thesis, optimal parameters to utilize GPUs to full scale were not available. Therefore, we have performed a series of test runs to determine the parameters to obtain optimal computational speed for our system.

Terminology [83]

- CPU- It stands for central processing unit and is often called the brains of a computer. It handles all computational work.
- GPU- It stands for graphics processing unit and is often used for rendering images of high quality. Recently, it is also used as a powerful computational device.
- Threads- A software thread is the smallest part of a program that can be processed independently by the operating system (OS) and a hardware thread is a physical CPU or a core that can process information. A single hardware thread can process

multiple software threads managed by the OS. In general, even a single CPU contains multiple cores and software manages their utilization.

- Core- A CPU can contain many hardware cores and each hardware core has an ability to process a set of instructions. These cores can contain one or more logical cores and each logical core has the ability to execute a single thread.
- CPU/GPU- The difference between a CPU and a GPU is in its architecture as shown in Fig []. A CPU has a fewer number of cores and can execute only few software threads compared to a GPU containing thousands of cores.
- Node- It is an object of network and the way different objects communicate. An example being communication between a computer and a printer. In an HPC, this refers to the collection of processors that can communicate with each other.
- MPI- It stands for message parsing interface and it provides an ability to utilize multiple processors.

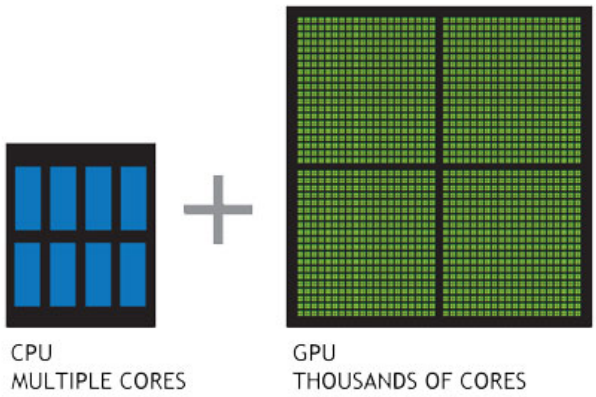


Fig A-1: Illustration showing the difference between the numbers of cores in a CPU and a GPU

In an MD simulation, the computationally expensive calculations are determining the short-range non-bonded interactions as long-range non-bonded interactions are usually taken care by advance algorithms such as, PME. In GROMACS [], these short-range calculations are off loaded to GPUs and the CPUs compute all-remaining forces such as bonded interactions, PME calculations and also update the particle positions.

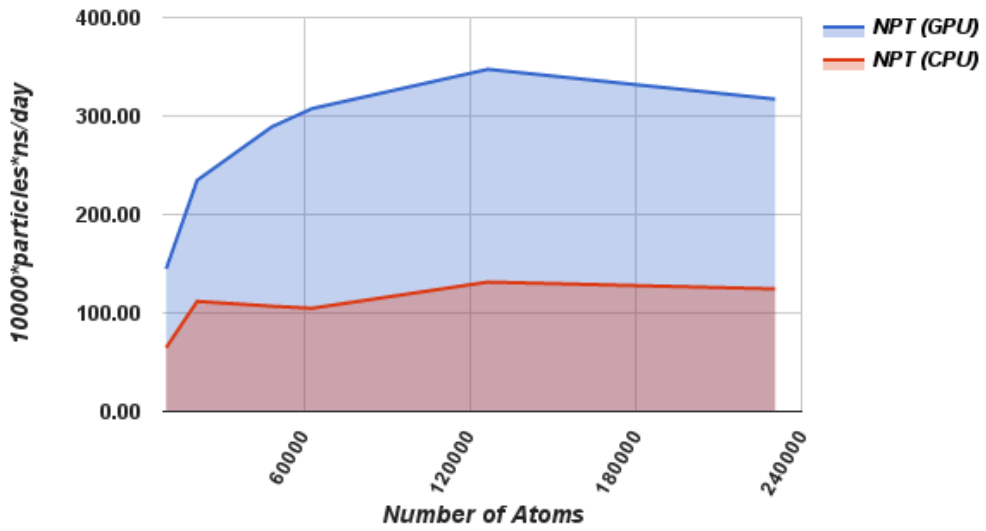


Figure A-2: Comparison of computational speed between CPU and GPU for differ number of particles. 20 CPUs and 2 GPUs of k20 model are used for this analysis.

. In order to utilize the maximum efficiency of the available hardware GROMACS uses thread MPIs for a single processor and OPEN MPIs for multiple processors. Therefore, the computational performance of the simulation depends on efficiency of communication between CPU and GPU as individually they process at their maximum

efficiency. mdrun command of GROMACS internally obtains the optimal parameters based on the available hardware from version 5.0. However, we have seen that if the hardware is not chosen carefully, there could be a significant loss in the computational performance. Fig A-2 shows the comparison of the simulation performance for different system sizes, where the y-axis is the speed in ns/day normalized by the number of particles in each system. All the simulations are performed using k20 model GPUs and a single node with 20 CPUs, 4 MPI processors, 2 GPUs (shared by the 4 MPIs) and all threads are pinned to the core manually by using -pinon option of GROMACS v 5.0.2. This shows that using GPUs the computational speed is enhanced by 3 times. We have also noticed a significant increase in performance when using thread MPI compared to OPEN MPI in a single node and decrease in performance when the –pinon option was not used. We make a note that, the performance analysis was not done on the same nodes and other programs running on the node can have an influence on the performance, but it does provide a good estimate.

In conclusion, we have performed molecular dynamic simulations with the above mentioned settings for our BSA-CNM system consisting around 240,000 atoms and ran for 100 ns in ~6 days real time.

Appendix B

SUPPLEMENTARY INFORMATION FOR CHAPTER 3

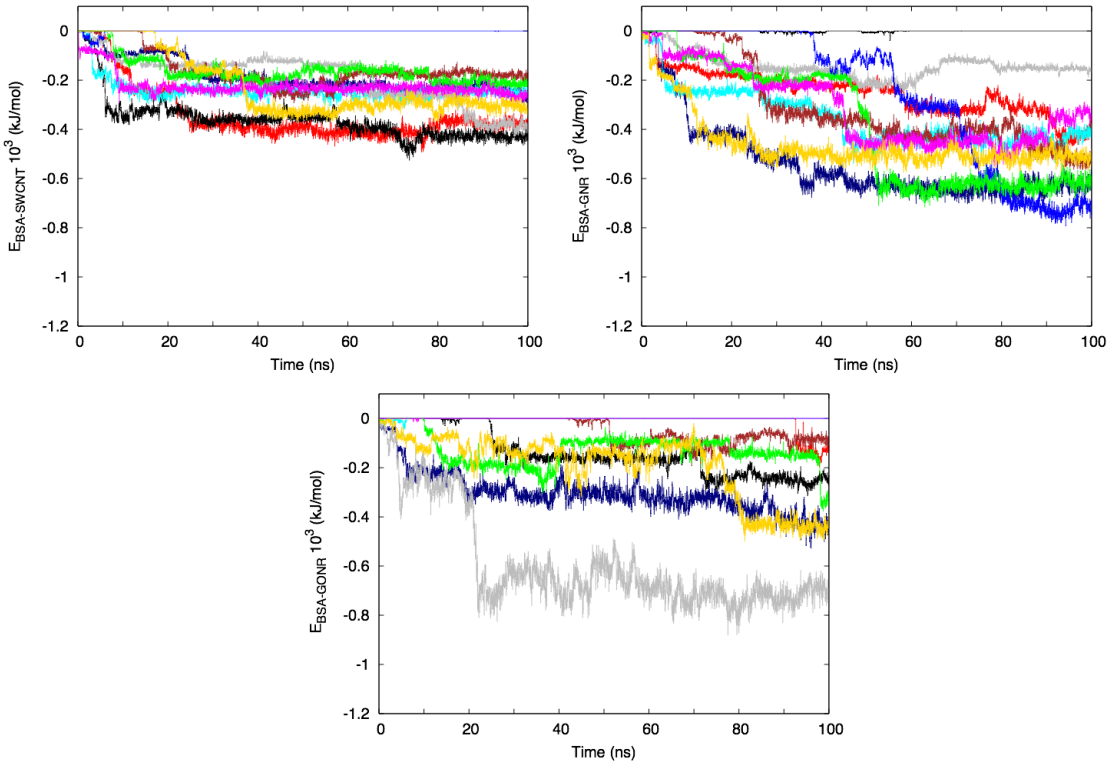


Fig B-1: Interaction Energies for all starting configurations for different CNMs. Notice that in some cases there is no interaction between the BSA and CNM.

`

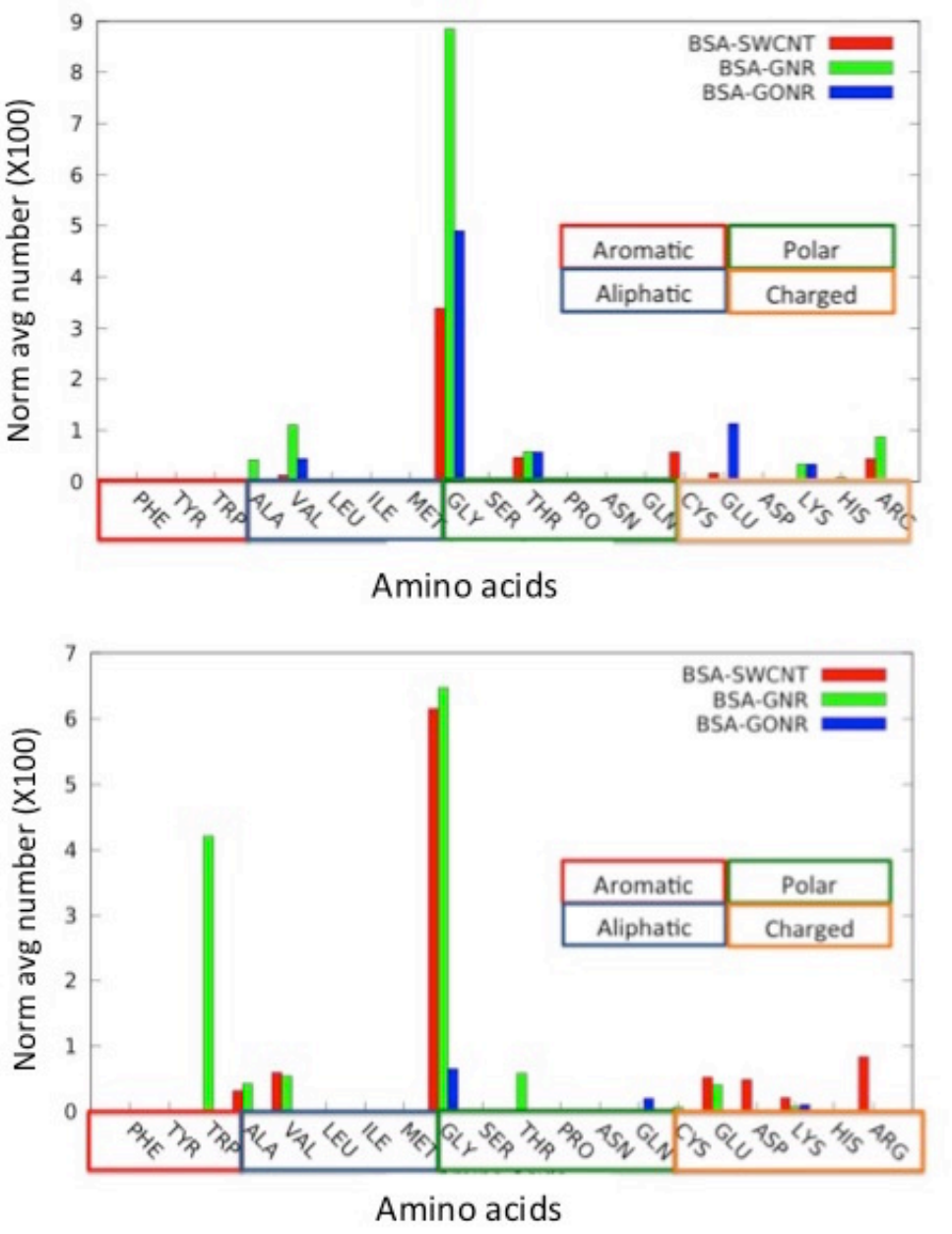


Figure B-2: Average adsorption number of residues in the interaction zone for different starting configurations.

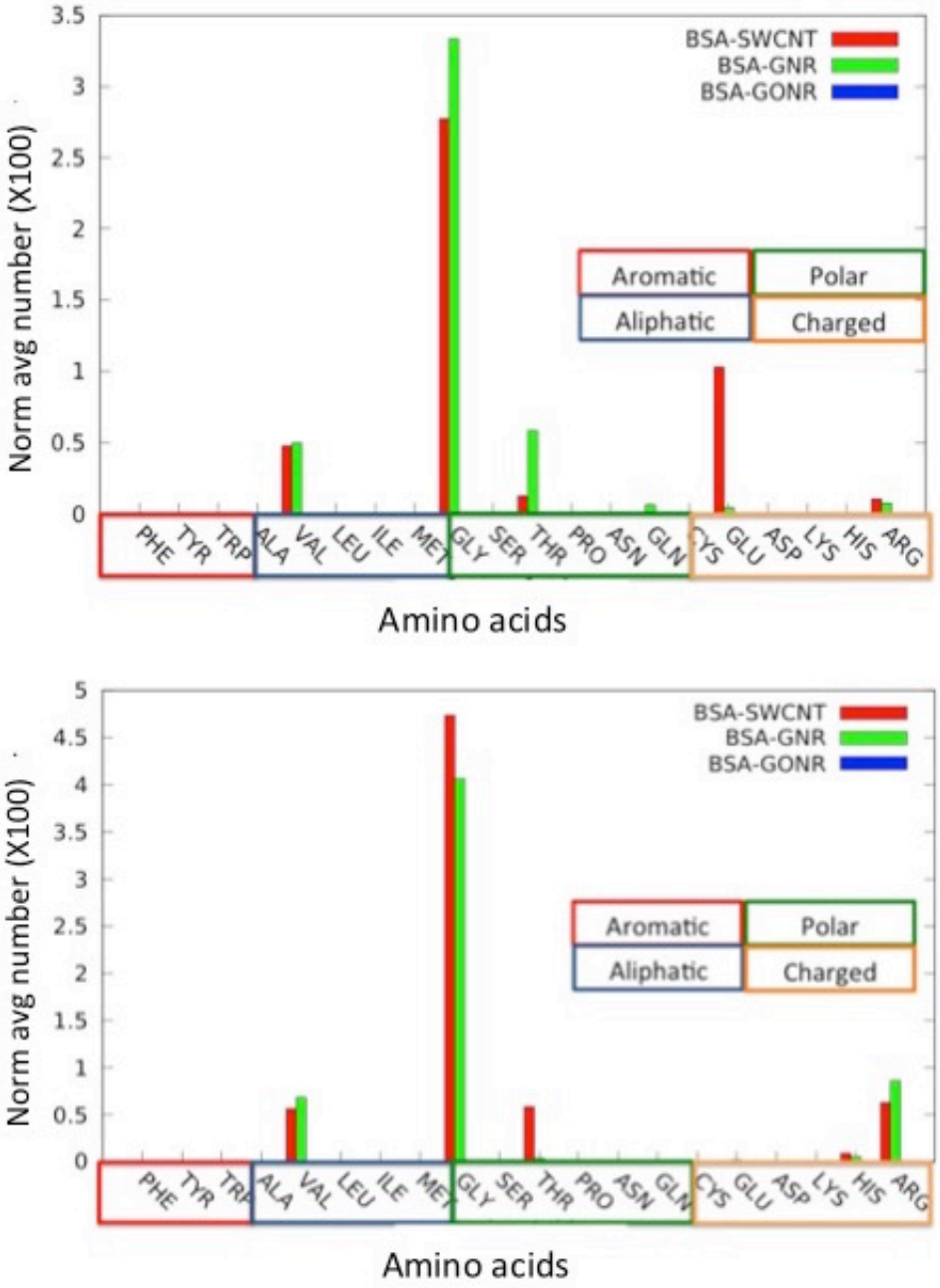
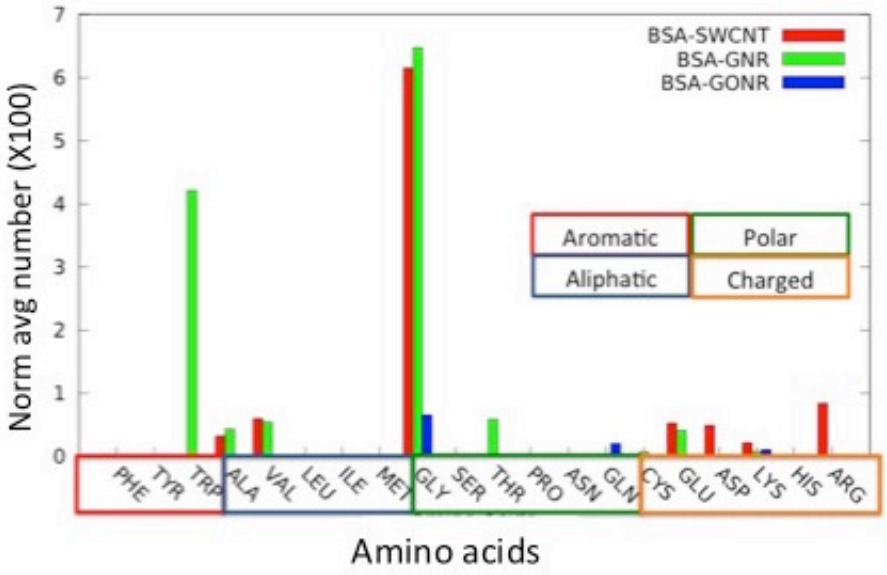


Figure B-2 (continued): Average adsorption number of residues in the interaction zone for different starting configurations.



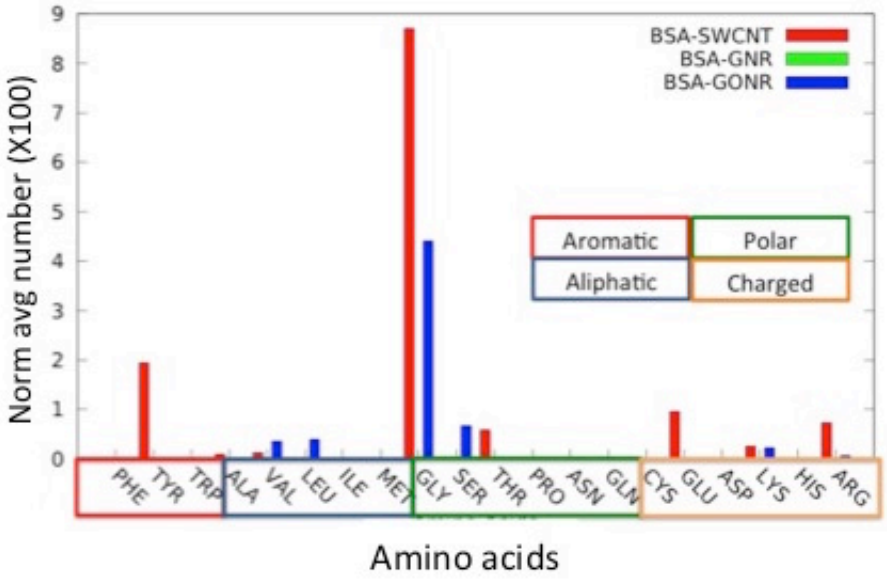


Figure B-2 (continued): Average adsorption number of residues in the interaction zone for different starting configurations.

Appendix C

TEMPERATURE COUPLING GROUPS AND TEMPERATURE DISTRIBUTIONS

We performed a set of simulations to analyze the problem of hot solvent and cold solute. This problem means, development of temperature gradients in the system. We found that using a) different temperature coupling groups for the solute and the solvent is effective than b) combining them into a single coupling group. However, we did not see any significant differences in the temperature by increasing the temperature coupling time for the solute. The test analysis is performed on a tripeptide-graphene system by fixing the temperature to 300 K using noose-hoover thermostat. Before running the simulations, we expected due to the small solute, combining both solute and solvent into a single coupling group might be effective, but the results are contrary. Table C1 shows the average temperatures in the two cases.

Table C1: Average temperatures of solute and solvent

Temperature	Case a (K)	Case b (K)
Peptide	537.88	300.091
Solvent	296.623	300.08

As seen, in the case b the average temperature of the peptide is much higher than case a. When the temperature fluctuations are plotted (Fig C1), we observe a significant increase in instantaneous temperatures as reflected in the average temperatures. This shows the ineffectiveness of Case a.

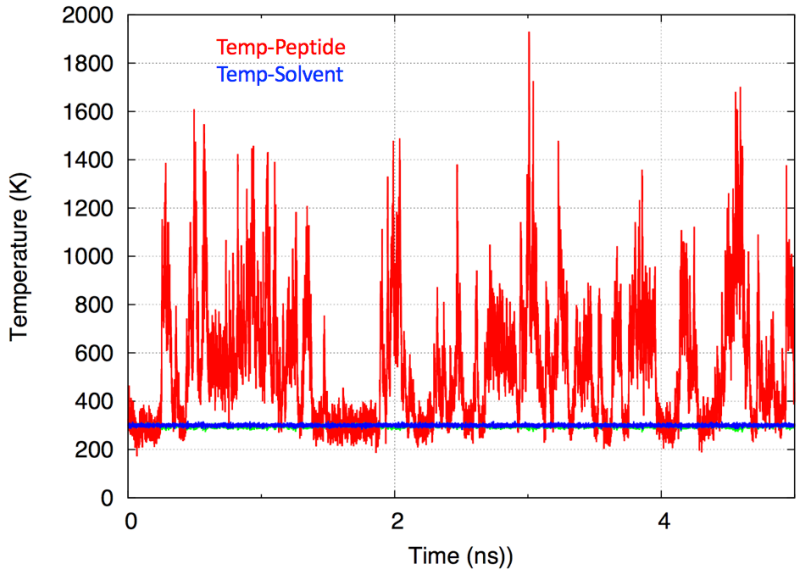


Figure C-1: Temperature fluctuations of the tripeptide-GRA-solvent system where the combined temperature-coupling group is used for both peptide and solvent.

In order to check the dependence of time coupling constant, we have also performed a set of simulations for the same system using case b by increasing it from 5 ps to 50 ps. In all the cases, we have observed the average temperatures and fluctuations are similar. We also plotted the temperature distributions in each case and fitted it to a chi-square distribution and did not observe any significant changes as shown in Fig C-2

when temperature coupling time is 5ps. Therefore, we have used case b in all our simulations performed in chapter 4 and used 5 ps as temperature coupling time for the peptide, which is in accord with Lingenheil et al. [84].

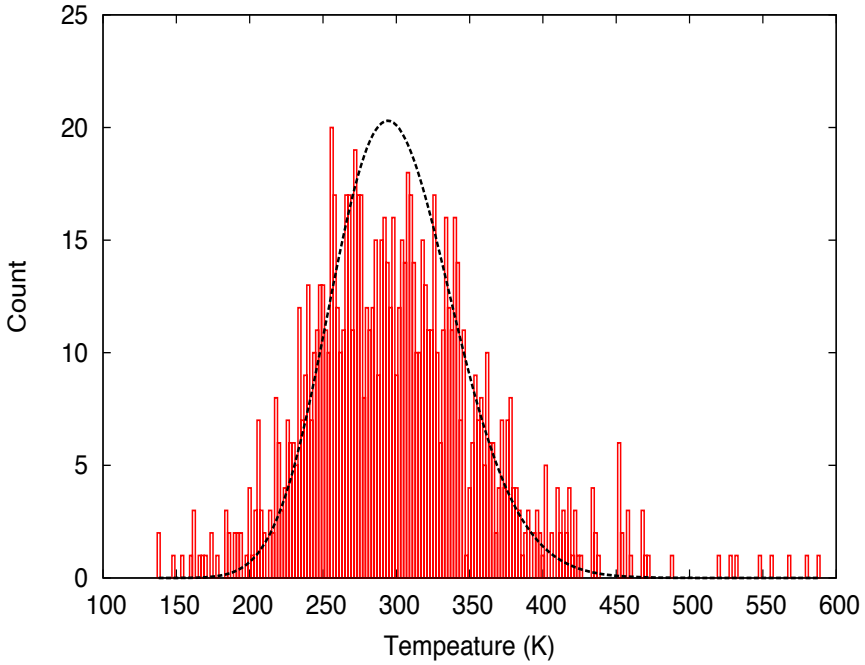


Figure C-2: Temperature distribution fitted to chi-square distribution when the temperature coupling time is 5 ps.

Appendix D

UMBRELLA SAMPLING USING GROMACS

We first validated our umbrella sampling simulations performed using pull code section in GROMACS v 5.0.2 [75] by calculating PMF of a methane-methane system and comparing the results with the PMF obtained from radius of gyration and the available literature data. Following is the procedure to run such a simulation and the validation of the results.

Procedure and Pull Code Parameters in GROMACS

- Choose a reaction coordinate for the system. We have used the distance between both the united atom methane atoms as the reaction coordinated as illustrated in Fig 4.1
- Set up the starting configurations for each window based on the chosen reaction coordinate. We froze one methane atom and placed the other at varying distances. Then, we solvated the box with tip3p model water molecules.
- Control umbrella sampling simulations for each window using the following set of parameters. The values listed correspond to our methane-methane system.
    - pull = umbrella (Turn on umbrella sampling)
    - pull-geometry = distance (Use distance as order parameter)
    - pull-dim = Y Y Y (distance measured using all components)
    - pull-start = yes (use starting position as reference position)

- pull-coord1-init = 0 (set this to 0)
- pull-ngroups = 2 (number of groups, solute and surface)
- pull-coord1-groups = 1 2 (identify each group using index file)
- pull_group1-name = MET1 (name of first group)
- pull_group2-name = MET2 (name of second group)
- pull_coord1-k = 3000 (force constant in kJ mol^-1 nm^-2)
- pull-nstxout = 500 (frequency of writing output)

- Run the simulations and determine PMF using g_wham tool available in GROMACS, which performs weighted histogram analysis to unweight the biased potential.

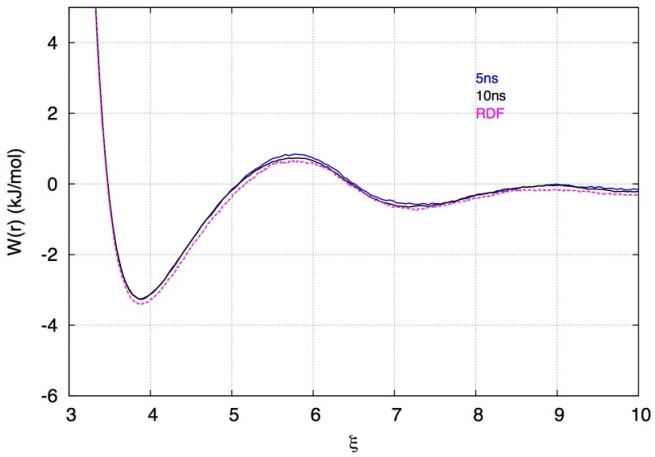


Figure D-1: PMF of a methane-methane system compared with PMF obtained from radial distribution function. Both 5ns and 10 ns have resulted in the same PMF indicating convergence. Note: A correction factor is applied according to Neumann et al. [85]

Appendix E

INTRODUCTION TO IMPLICIT SOLVENT[86]

Implicit Solvent- In addition to an explicit solvent model, molecular dynamics simulations can also use an implicit solvent model. In an implicit solvent, the solvent is modeled as a continuum, which approximates the behavior of all water molecules. These models are developed by breaking down the interactions between solvent and solute into a) non polar and b) electrostatic interactions. In general, various methods were developed to model the electrostatic interactions, as they are the computationally expensive part of the interactions. In order to model these interactions, different methods have been proposed such as,

- Accessible surface area based method (ASSA)
- Poisson Boltzmann method (PB)
- Generalized Born method (GB)
- Hybrid implicit-explicit solvent methods.

The following is a brief introduction to the first three methods.

ASSA- This is the simplest method where both the interactions are approximated by determining the solvent accessible surface area and a solution parameter of each atom that approximates the free energy of solvation (energy required to transfer a solute from vacuum to solvent). However, this method is used widely to calculate only the non-polar interactions.

PB- In this method, all non polar interactions are usually modeled by ASSA and the electrostatic interactions are modeled by using the charge density of the solute and electrostatic potential is determined using the Poisson- Boltzmann (PB) equation given by,

$$\vec{\nabla}.[\varepsilon(\vec{r})\vec{\nabla}\psi(\vec{r})] = -4\pi\rho^{f}(\vec{r}) - 4\pi\sum_{i} c_i^{\infty} z_i q \lambda(\vec{r}) e^{-\frac{z_i q \psi(\vec{r})}{kT}}$$

Where, the first term on the right hand side describes the charge density of the solute and the second term is modeled to capture the effects due to the presence of ions. The solution to this equation depends on the accuracy of determining charge density. As the equation is highly non linear, it is solved by employing advanced numerical methods. Apart from the complex calculations, encoding this to the existing molecular dynamics algorithms is not an easy task and only few MD packages support this method.

GB- This is the most widely implemented method in most of the MD packages because of its simplicity. It is an approximation of PB equation and is often validated by the comparing the results of this method with the PB method. In this method, the solute is modeled as spheres of low dielectric constant surrounded by a large dielectric medium (solvent) and the solvation energy are computed by the following equations,

$$G_s = \frac{1}{8\pi}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon}\right)\sum_{i,j}^{N}\frac{q_i q_j}{f_{GB}}$$

where,

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}}$$

and

$$D = \left(\frac{r_{ij}}{2a_{ij}}\right)^2, \quad a_{ij} = \sqrt{a_i a_j}$$

In order to implement this method, the born radius of the atom (a_i) should be determined accurately. The advantage of this method is its simpler form compared to non-linear PB. Recently various algorithms (HCT, OBC and still) have been developed by different research groups to implement this method in all MD packages. In some of these algorithms, the non-polar interactions are calculated by using born radius which enhances the computationally speed compared to algorithms calculating it by using ASSA. However, this method is also computationally expensive as it is necessary to calculate the value of (a_i) for every step in MD simulations and is usually implanted using GPUs. The idea behind the last method (hybrid implicit-explicit) is to model explicit solvent for the water molecules near the solute and the rest by using an implicit solvent model.

We have explored the computational speeds in various MD packages and discovered that the reported benchmarks are often based on the simulations, where the implicit solvent includes only the electrostatic interactions. Since the bottle neck of implicit solvent algorithms is often the communication between CPUs and GPUs [82], we suggest using Amber MD package which utilizes only GPUs. However, due to the significant loss in computational performance with the current algorithms when the non-polar interactions are included we prefer explicit solvent to implicit solvent. The advantage of this method is that it enhances the sampling of the energy landscape of the solute.

REFERENCES

1. Buzea, C., I.I.P. Blandino, and K. Robbie, *Nanomaterials and nanoparticles: Sources and toxicity.* Biointerphases, 2007. **2**(4): p. MR17-MR172.
2. Eadie, L. and T.K. Ghosh, *Biomimicry in textiles: past, present and potential. An overview.* J R Soc Interface, 2011. **8**(59): p. 761-75.
3. Ali, M.E., et al., *Gold Nanoparticle Sensor for the Visual Detection of Pork Adulteration in Meatball Formulation.* Journal of Nanomaterials, 2012. **2012**: p. 1-7.
4. Li, W.R., et al., *Antibacterial activity and mechanism of silver nanoparticles on Escherichia coli.* Appl Microbiol Biotechnol, 2010. **85**(4): p. 1115-22.
5. Miller, G. and L. Sales, *Nano-Ingredients in sunscreen: The need for regulation.* Public Health Association, Australia, 2012.
6. Bianco, A., K. Kostarelos, and M. Prato, *Applications of carbon nanotubes in drug delivery.* Curr Opin Chem Biol, 2005. **9**(6): p. 674-9.
7. Ajayan, P.M. and O.Z. Zhou, *Applications of Carbon Nanotubes.* Carbon Nanotubes, Topics Appl. Phys., 2001. **80**: p. 391-425.
8. Tsuji, J.S., et al., *Research strategies for safety evaluation of nanomaterials, part IV: risk assessment of nanoparticles.* Toxicol Sci, 2006. **89**(1): p. 42-50.
9. Deng, J., D. Yu, and C. Gao, *Biological identity of nanomaterials: Opportunities and challenges.* Science China Chemistry, 2013. **56**(11): p. 1533-1541.
10. Library, S.P., http://www.sciencephoto.com.

11. Presswire, M., *Global Nanotechnology Market Outlook 2015-2020 - Industry will grow to reach US$ 75.8 Billion.* Business Insights: Essentials. , 11 June 2015. .

12. Feynman, R.P., *Plenty of Room at the Bottom.* 1959.

13. Michel, J.-B., et al., *Quantitative Analysis of Culture Using Millions of Digitized Books.* Science, 2010.

14. http://www.nature.com/subjects/nanotoxicology, *nature.com subject areas.*

15. Goldstick, M., *The Ability of Alpha Radiation To Penetrate Human Skin.* WISE, 1992.

16. Yah, C.S., G.S. Simate, and S.E. Iyuke, *Nanoparticles toxicity and their routes of exposures.* Pak. J. Pharm. Sci. **25**: p. 477-491.

17. Iijima, S., *Helical microtubules of graphitic carbon.* Letters to Nature, 1991.

18. Novoselov, K.S., et al., *Electric Field Effect in Atomically Thin Carbon Films.* Science, 2004. **306**.

19. Kashiwada, S., *Distribution of Nanoparticles in the See-through Medaka (Oryzias latipes).* Environmental Health Perspectives, 2006.

20. Poland, C.A., et al., *Carbon nanotubes introduced into the abdominal cavity of mice show asbestos-like pathogenicity in a pilot study.* Nat Nanotechnol, 2008. **3**(7): p. 423-8.

21. Jacobs, M.M., et al., *Precarious Promise: A Case Study of Engineered Carbon Nanotubes.* 2014.

22. Madani, S.Y., A. Mandel, and A.M. Seifalian, *A concise review of carbon nanotube's toxicology.* Nano Rev, 2013. **4**.

23. Wang, K., et al., *Biocompatibility of Graphene Oxide.* Nanoscale Research Letters, 2010.

24. Salvador-Morales, C., et al., *Binding of pulmonary surfactant proteins to carbon nanotubes; potential for damage to lung immune defense mechanisms.* Carbon, 2007. **45**(3): p. 607-617.

25. Karajanagi, S.S., et al., *Protein-Assisted Solubilization of Single Walled Carbon Nanotube.* Langmuir, 2006. **22**: p. 1392-1395.

26. Ge, C., et al., *Binding of blood proteins to carbon nanotubes reduces cytotoxicity.* Proc Natl Acad Sci U S A, 2011. **108**(41): p. 16968-73.

27. Raffaini, G. and F. Ganazzoli, *Surface topography effects in protein adsorption on nanostructured carbon allotropes.* Langmuir, 2013. **29**(15): p. 4883-93.

28. Podila, R. and J.M. Brown, *Toxicity of Engineered Nanomaterials: A Physicochemical Perspective.* J BIOCHEM MOLECULAR TOXICOLOGY, 2013. **27**.

29. Nel, A.E., et al., *Understanding biophysicochemical interactions at the nano-bio interface.* Nat Mater, 2009. **8**(7): p. 543-57.

30. Jimenez-Cruz, C.A., S.-g. Kang, and R. Zhou, *Large scale molecular simulations of nanotoxicity.* WIREs Syst Biol Med, 2014. **6**: p. 329-343.

31. Zhang, Y., et al., *Interactions of graphene and graphene oxide with proteins and peptides.* Nanotechnology Reviews, 2013. **2**(1).

32. Zuo, G., et al., *Interactions between proteins and carbon-based nanoparticles: exploring the origin of nanotoxicity at the molecular level.* Small, 2013. **9**(9-10): p. 1546-56.

33. Chen, J., et al., *Molecular dynamics simulations of the adsorption of DNA segments onto graphene oxide.* Journal of Physics D: Applied Physics, 2014. **47**(50): p. 505401.

34. Zuo, G., et al., *Plugging into Proteins: Poisoning Protein Function by a Hydrophobic Nanoparticle.* ACS Nano, 2010. **4**.

35. Zuo, G., et al., *Adsorption of Villin Headpiece onto Graphene, Carbon Nanotube, and C60: Effect of Contacting Surface Curvatures on Binding Affinity.* The Journal of Physical Chemistry C, 2011. **115**(47): p. 23323-23328.

36. Wang, Y. and H. Ai, *Theoretical Insights into the Interaction Mechanism between Proteins and SWCNTs: Adsorptions of Tripeptides GXG on SWCNTs.* J. Phys. Chem, 2009. **113**: p. 9620-9627.

37. Wang, S., et al., *Peptides with selective affinity for carbon nanotubes.* Nat Mater, 2003. **2**(3): p. 196-200.

38. Xie, H., et al., *Ranking the affinity of aromatic residues for carbon nanotubes by using designed surfactant peptides.* J Pept Sci, 2008. **14**(2): p. 139-51.

39. Rajesh, C., et al., *A theoretical study on the interaction of aromatic amino acids with graphene and single walled carbon nanotube.* J Chem Phys, 2009. **130**(12): p. 124911.

40. He, Z. and J. Zhou, *Probing carbon nanotube–amino acid interactions in aqueous solution with molecular dynamics simulations.* Carbon, 2014. **78**: p. 500-509.

41. Camden, A.N., S.A. Barr, and R.J. Berry, *Simulations of peptide-graphene interactions in explicit water.* J Phys Chem B, 2013. **117**(37): p. 10691-7.

42. Hughes, Z.E. and T.R. Walsh, *What makes a good graphene-binding peptide? Adsorption of amino acids and peptides at aqueous graphene interfaces.* J. Mater. Chem. B, 2015. **3**(16): p. 3211-3221.

43. Zhang, M., et al., *Interaction of peptides with graphene oxide and its application for real-time monitoring of protease activity.* Chem Commun (Camb), 2011. **47**(8): p. 2399-401.

44. Gu, Z., et al., *The role of basic residues in the adsorption of blood proteins onto the graphene surface.* Sci Rep, 2015. **5**: p. 10873.

45. Wu, E., M.O. Coppens, and S. Garde, *Role of arginine in mediating protein-carbon nanotube interactions.* Langmuir, 2015. **31**(5): p. 1683-92.

46. Anversa, J. and P. Piquini, *The effects of an explicit water environment on the interaction of a single wall carbon nanotube with amino acids: A theoretical study.* Chemical Physics Letters, 2011. **518**: p. 81-86.

47. Lu, C.H., et al., *General approach for monitoring peptide-protein interactions based on graphene-peptide complex.* Anal Chem, 2011. **83**(19): p. 7276-82.

48. Shen, J.W., et al., *Adsorption of insulin peptide on charged single-walled carbon nanotubes: significant role of ordered water molecules.* Chemphyschem, 2009. **10**(8): p. 1260-9.

49. Penna, M.J., et al., *Molecular-level understanding of the adsorption mechanism of a graphite-binding peptide at the water/graphite interface.* Soft Matter, 2015. **11**(26): p. 5192-203.

50. Raffaini, G. and F. Ganazzoli, *Molecular Dynamics Simulation of the Adsorption of a Fibronectin Module on a Graphite Surface.* Langmuir, 2004. **20**: p. 3371-3378.

51. Raffaini, G. and F. Ganazzoli, *Simulation Study of the Interaction of Some Albumin Subdomains with a Flat Graphite Surface.* Langmuir, 2003. **19**: p. 3403-3412.

52. Mijajlovic, M., M.J. Penna, and M.J. Biggs, *Free energy of adsorption for a peptide at a liquid/solid interface via nonequilibrium molecular dynamics.* Langmuir, 2013. **29**(9): p. 2919-26.

53. Raffaini, G. and F. Ganazzoli, *Protein adsorption on a hydrophobic surface: a molecular dynamics study of lysozyme on graphite.* Langmuir, 2010. **26**(8): p. 5679-89.

54. Balamurugan, K., et al., *Exploring the Changes in the Structure of r-Helical Peptides Adsorbed onto a Single Walled Carbon Nanotube Using Classical Molecular Dynamics Simulation.* J. Phys, 2010. **114**: p. 14048-14058.

55. Mucksch, C. and H.M. Urbassek, *Molecular dynamics simulation of free and forced BSA adsorption on a hydrophobic graphite surface.* Langmuir, 2011. **27**(21): p. 12938-43.

56. Mücksch, C. and H.M. Urbassek, *Adsorption of BMP-2 on a hydrophobic graphite surface: A molecular dynamics study.* Chemical Physics Letters, 2011. **510**(4-6): p. 252-256.

57. Balamurugan, K. and V. Subramanian, *Length-dependent stability of alpha-helical peptide upon adsorption to single-walled carbon nanotube.* Biopolymers, 2013. **99**(6): p. 357-69.

58. Karunwi, O., et al., *Molecular Dynamics Simulations of Peptide–Swcnt Interactions Related to Enzyme Conjugates for Biosensors and Biofuel Cells.* Nano LIFE, 2013. **03**(04): p. 1343007.

59. Shen, J.W., et al., *Induced stepwise conformational change of human serum albumin on carbon nanotube surfaces.* Biomaterials, 2008. **29**(28): p. 3847-55.

60. Gu, Z., et al., *Surface Curvature Relation to Protein Adsorption for Carbon-based Nanomaterials.* Scientific Reports, 2015.

61. Walsh, T.R. and S.M. Tomasio, *Investigation of the influence of surface defects on peptide adsorption onto carbon nanotubes.* Mol Biosyst, 2010. **6**(9): p. 1707-18.

62. Sengupta, B., et al., *Influence of carbon nanomaterials defects on the formation of protein corona.* submitted to PCCP, 2015.

63. Peng, X., et al., *Adsorption of human serum albumin onto highly orientated pyrolytic graphite surface studied by atomic force microscopy.* Scanning, 2015. **37**(2): p. 158-64.

64. News, S., *Grand View Research Anticipates Carbon Nanotubes (CNT) Market To Grow To $3.42 Billion By 2022.* http://www.scopii.com/news-collection/2015/05/29/grand-view-research-anticipates-carbon-nanotubes-cnt-market-to-grow-to-3-42-billion-by-2022/, 2015.

65. Bujacz, A., *Structures of bovine, equine and leporine serum albumin.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 10): p. 1278-89.

66. Cornell, W.D., et al., *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules.* J Am Chem Soc, 1995. **117**: p. 5179-5197.

67. Abraham, M.J., et al., *GROMACS User Manual version 5.0.* 2014.

68. Berman, H.M., et al., *The Protein Data Bank* Nucleic Acids Research, 2000. **28**: p. 235-242.

69. DeFever, R.S., et al., *PAMAM Dendrimers and Graphene: Materials for Removing Aromatic Contaminants from Water.* Environ. Sci. Technol., 2015. **7**: p. 4490-4497.

70. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926.

71. Lindorff-Larsen, K., et al., *Improved side-chain torsion potentials for the Amber ff99SB protein force field.* Proteins, 2010. **78**(8): p. 1950-8.

72. G, B., D. D, and P. M, *Canonical sampling through velocity rescaling.* J Chem Phys. , 2007.

73. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath* J. Chem. Phys. , 1984. **84**.

74. Essmann, U., et al., *A smooth particle mesh Ewald method.* J. Chem. Phys., 1995. **103**.

75. Hess, B., et al., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.* J. Chem. Theory Comput, 2008. **4**.

76. Berk Hess, H.B., Herman J. C. Berendsen andJohannes G. E. M. Fraaije, *LINCS: A linear constraint solver for molecular simulations.* J. Comput. Chem., 1998. **18**: p. 1463–1472.

77. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features.* Biopolymers, 1983. **22**(12): p. 2577-2637.

78. Shell, M.S., *Biased sampling and related free energy techniques.* Lecture Notes, 2009.

79. Marc Souaille, B.t.R., *Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations.* Computer Physics Communications 2001. **135** p. 40–57.

80. Patel, A.J., et al., *Quantifying density fluctuations in volumes of all shapes and sizes using indirect umbrella sampling.* Journal of Statistical Physics 2011.

81. http://citi.clemson.edu/palmetto/pages/userguide.html.

82. Kutzner, C., et al., *Best bang for your buck: GPU nodes for GROMACS biomolecular simulations.* 2015.

83. http://www.howtogeek.com/194756/cpu-basics-multiple-cpus-cores-and-hyper-threading-explained/.

84. Lingenheil, M., et al., *The "Hot-Solvent/Cold-Solute" Problem Revisited.* JCTC, 2008. **4**(1293-1306).

85. Neumann, R.M., *Entropic approach to Brownian movement.* American Journal of Physics, 1980. **48**(5): p. 354.

86. Onufriev, A., *The generalized Born model: its foundation, applications, and limitations.* Review Notes, 2010.